Legislative voting and party unity in the European Parliament

by

Jonas Nordkvelle

Department of Political Science
Faculty of Social Sciences
University of Oslo
May 2012

Legislative voting and party unity in the European Parliament.

`http://www.duo.uio.no/`

Print: Reprosentralen, University of Oslo

Legislative voting and party unity in the European Parliament

Abstract

Why do legislators in some political parties choose to vote coherently over time and over a varied set of policy proposals, while less so in others? Are there structural differences between these parties that increase or decrease the vote cohesiveness? And are there specific situations where the cohesiveness breaks down? Investigating party unity in legislatures is important, because it can tell the electorate which groups are relevant actors within the legislature, and what outcomes to expect if voting for legislator A or party B – information that can increase the democratic accountability. In addition, discovering mechanisms that increase/decrease party unity is interesting for political scientists and political entrepreneurs, because it can help discover problems of democratic accountability and help propose solutions to improve it.

Previous research have proposed four sources for observed party unity in legislatures: preferences, party discipline, agenda-setting and division-of-labour. I link these sources to specific hypotheses regarding the effect of situational and structural variables, which I put to test using multi-level logit regression models on roll-call data from the 2009-10 European Parliament. The massive amount of national parties and their affiliation with party groups, as well as variation on electoral institutions, make the EP a haven for political research.

I have two main contributions. First, I argue that defection by larger groups (like the majority in a national party) shows a different pattern than defection by individuals. This is very relevant when it comes to researching the European Parliament, but also legislatures with party coalitions. Second, I have coded candidate selection methods for most of the parties represented in the EP, and tested the claim that centralized candidate selection increases the ability of party leadership to dicipline their representatives, thus reducing the probability of defection from the national party. Contrary to what was found in Hix (2004), differences in candidate selection methods are not found to have the hypothesized effect. It may be that national party cohesion now is so high, that the proposed effect of candidate selection is less relevant.

Acknowledgments

This thesis could not have been made without the excellent teaching, guidance and advise of my supervisor, Bjørn Høyland. Without him, I would not have had the skills nor knowledge needed to write this thesis.

I am grateful to Espen Geelmuyden Rød not only for reading my paper and giving me helpful comments, but for always having a positive spirit and yearning for learning. Working beside him, and with him, has been a huge learning experience and great fun.

I am grateful to Ingrid, who graciously asked me if I wanted her to read my paper, and for the comments she gave me when I said yes. Thanks to her, this paper will be much easier to read.

I am grateful to all you nice people who were living your lives on the ninth floor with me, for making coffee-breaks long and entertaining.

I am grateful to friends and family for support and for making my life a good one.

Most of all, I am grateful to Annette for her love and support, as well as for reading thoroughly through the paper and giving comments.

The responsibility for any omissions, errors or mistakes are mine.

Jonas Nordkvelle
Oslo, 16 May 2012
Word count: 30699

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In this paper, I investigate why legislators from time to time choose to vote differently from the majority in their party (*'a vote defection'*). In the US, this phenomenon is so common that one is talking about degree of party cohesion – not bothering about the individual event. Individual legislators are placed in a policy-dimension (liberal-conservative) instead, showing that the most liberal Republican is just as liberal as the most conservative Democrat (Clinton, Jackman and Rivers, 2004). In Europe, with its tradition for strong parties, the norm is to vote the same as the rest of the party (Carey, 2009). That way, vote defection is a rare event, and becomes noteworthy in itself.

Research Question 1 *What explains vote defection in legislatures?*

More specifically, I will investigate the 2009 European Parliament (EP) election and the roll-call votes[1] collected in the period from September 2009 to October 2010. The European Parliament (EP) is a fascinating institution. The 754 Members of the EP (MEPs) are elected by 82 different electorates and they represent more than 500 million people in 27 different countries having (at least) 27 different public spaces. The representatives are organised in 172 national parties. Many of these are among the largest and most powerful organisations in the European countries, represented in the national legislature and (for about half of them) in national government. The national parties are again organised into 8 European Party Groups (EPGs) based on ideology. Christian Democrats, Social Democrats, Liberals, Conservatives and Greens all got their own group.

Research Question 2 *What explains vote defection in the 7th European Parliament?*

[1]Roll-call votes must be specifically called for by either a political group or at least 37 MEPs. Else, the vote is done by hand or electronically without individual votes being recorded (Hix and Høyland, 2011, p.60).

The main influencing article for this paper is an article written by Simon Hix in World Politics in 2004. My paper can be read as a replication study of that article for 10 year later data. One main point in that article is that the EP is especially interesting to do research on because it provides variation where other legislatures do not. The huge number of national parties, and the different electoral rules in each member state, make statistics about the effect of electoral institutions and the candidate selection process within each party on the probability of party defection possible.

1.1 Accountability

A main reason why we should investigate legislative voting is accountability (Carey, 2009). With accountability, Carey means "that legislators are responsive to the preferences and demands of their principal(s), that information about legislators' actions is available to the principal(s), and that principals can punish legislators for lack of responsiveness" (Carey, 2009, p.3). The typical principals for a legislator are the electorate and the party leadership, but others are also possible, like a president or a party group.

Whether the legislators are accountable to these principals as a group or as individuals depends on the information and tools of punishment/reward that are available to the principals. For instance, in a proportional representation (PR) vote system with closed lists, a large district magnitude, elite controlled candidate selection and no public information about what the individual legislators voted, legislators are individually accountable by their party leaderships, while they are only collective accountable by their electorate. The electorate has no way to punish individual legislators without also punishing the rest of that legislator's party, nor have they any way of keeping a record of how individual legislators are performing when voting, further reducing their control over individual representatives.

Such a system leads to strong (cohesive/unitary) parties, because the legislators have greater incentives to respond to the preferences and demands of the party leadership rather than to their electorate's preferences and demands. It follows from this that individual legislators are bound by their parties, and would-be-rebels get curbed. This is argued to be positive for these types of systems, because it implies that the parties become more united, and therefore more accountable as a group. Carey links this to what he calls 'decisiveness problems' (Carey, 2009, p.4f). Although the electorate would get more specific control over the legislators with individual accountability, no single legislator has the power to decide policy outcomes. It is much harder to hold individual legislators accountable for

policy outcomes. With strong parties, voters can hold large groups accountable for their actions (Bowler, Farrell and Katz, 1999, p.3).

However, if change is needed/wanted within parties, rather than change of party composition, strong parties and lack of individual accountability becomes a problem. Carey (2009) describes such problems arising in the Americas, and the reforms applied to correct it. Therefore, it is not a given that a highly cohesive legislature is a good sign. It could be a sign of exclusive elite control over the legislature, and a total lack of accountability to the electorate.

One of the main points with the EP is to provide the EU with a more directly democratic accountable political institution. The question then is whether the electorates in the EU have the information and tools available in order to make the legislators accountable to them. One giant effort made to give the electorates the information they need is the votewatch.eu project (Hix et al., 2012). They record all roll-call votes cast in the EU, and lobby for transparency within the EU. However, this information needs to be processed in order to produce relevant knowledge for the public.

For the electorate, being able to find out what the legislators voted is perhaps the most relevant information in this regard. They can also find out when groups that normally stand together break up - and maybe even some probable reasons for the break-up. It is interesting for the public to know what the group they are voting for are agreeing on and what they are not agreeing on. For instance, if the reason you voted for A is that A1 said she was against further integration of the EU, then it is of interest to know that group A does not seem to agree on this question. More subtly, you can find out how strong the group-bonds are by finding those situations where it might be tempting for some to vote differently. If a group is found to be very cohesive, it makes little sense to vote for the group because of what a legislator from the fringe of this group is saying she wants to do. Most probably, she will never be able to convince the rest of her group, and also – because the group is strong, she will probably vote the same as her group. More generally, the question we should pose is why a given group of legislators are voting together. While the answer to this could be straightforward – 'because they have the same preferences', this answer is not always true, and does not paint the whole picture. For instance, strategic voting – to vote something you do not agree with at a point in order to get in a better position for another vote (or something else), is probably common in modern legislatures. The bottom line is that an investigation of roll-call votes in terms of party unity/vote defection can give us some indications on answers to these questions – answers which are important for the electorate to know in order to make the legislators

sufficiently accountable to them.

1.2 Summary

In order to investigate my research question, I use a statistical analysis of roll-call data. I find that defection by national parties from their EPGs has decreased three-fold from when Hix did his analysis with data from 2000. Defection from national parties, on the other hand, has been relatively stable. This conforms with the observations that the EPGs have consolidated during the first decade in this millennium (Hix, Noury and Roland, 2009).

There are some differences between the EPGs in this respect. The Greens, S&D and EPP have very few defections, while ALDE and ECR (and the other smaller groups) have a bit more, but still few. ALDE probably have a bit more defections because they are placed in the middle of the political spectrum, and are some times able to tip the vote. ECR, on the other hand, is more likely to have real preferential differences between the national parties affiliated with it. This is not surprising, considering that it is a newly founded group.

Electoral institutions were found to have an effect on party unity in Hix (2004) and Faas (2003). I also find effects that are consistent with my data, but they have wrong sign – being positive when the theory predicts negative, and vice versa, for all effects that have to do with candidate selection. This result is curious, and can have many different explanations, as I will elaborate on in the concluding chapter.

The high cohesion we see in the European party groups has been interpreted by others to show that the European Parliament is starting to function as a real parliament, and it being an indicator for that the EP is becoming a more relevant political institution with real power (Hix, Noury and Roland, 2007). A consequence of this real power is that it becomes more important that its members are accountable to the electorate. As long as the electorates, national media and national party members believe that it is the national parties that have power in the European Parliament – and just as much as they have in the national parliament, and not constrained by the party group they are affiliated with, then the EP has a problem with accountability. The EPGs show all signs of functioning as parties in a national strong party legislature, while the national parties seem to work more like single legislators in a national strong party legislature. Within all such legislatures, the electorate votes for the parties and not for single legislators (or if they do, they at least know which party that legislator is affiliated with). On this background, I argue that in order to improve accountability, national parties need to go to election under their

European party group banner.

1.3 Structure of paper

Chapter 2: Here I assemble a general theory of legislative voting, and indicate possible sources that can increase observed party unity. Carey (2009) and Hix, Noury and Roland (2009) propose that the four most important sources of party unity are preferences, discipline, agenda control and division-of-labour. I elaborate on these four sources.

Chapter 3: The 2004 article by Simon Hix is central to this paper, and I have devoted this chapter to it. In the article, Hix produces a theory of his own, which I have only loosely followed. However, the ideas about electoral institutions and how they are related to party discipline are more or less blue-printed in this paper. Furthermore, the use of party defection as dependent variable, and the following test using logit models, are followed in my paper too. Hix' central claim is that electoral institutions, including candidate selection methods, "shape the relationship between elected politicians and their local and parliamentary principals" through giving/taking power to discipline the legislators (Hix, 2004, p.219).

One of the contentions I have with the 2004 article is that the operationalisation of candidate selection is cursory. I will therefore investigate candidate selection in the light of an excellent book by Hazan and Rahat (2010) and discuss ways to operationalise candidate selection in a proper and useful way. The coding of candidate selection methods for national parties before the 2009 EP election has been one of the major undertakings in this paper, as there was no up to date or comprehensive data on candidate selection methods in European parties for the EP elections when I started. Fortunately, an excellent report on the subject commissioned by the Committee on Constitutional Affairs in the EP before the 2009 EP election (Lehmann, 2009) was available. Because of that report, I have been able to give meaningful codes to a large majority of the parties represented in the EP.

Chapter 4: As I have already shown, party unity can be thought of in different ways. In Chapter 4, I will review these and provide reasons for why I choose to use party defection as the dependent variable in the regression models. I define two types of party defection in the EP: defection by the national party from the EPG, and defection by individual MEPs from the national party. The reason for doing so, is that these types of defections have different explanations.

Chapter 5: Chapter 5 will provide the hypotheses I will be testing. There is a huge discrepancy between what we would like to test given our theories and what kind of data we actually have available. However, while most of the hypotheses have been used before in other settings, some are new – hopefully contributing to the research. The hypotheses relate to type of defection and the sources of party unity theorized in Chapter 2.

Chapter 6: Chapter 6 presents the data and regression models that I will use to test the hypotheses. The methods I will use in order to test the hypotheses are a mix of bivariate graphical analysis and pooled, unpooled (fixed-effects) and partially pooled (multi-level) regression analysis.

Chapter 7: Here, the results of the regression models will be presented and analysed. I will also compare some of the models and show the predictive power of some of them. Finally, I will shortly investigate the use of the 'Abstain'-option when voting, which can say something about how MEPs and parties choose to defect. The reason for this is that voting 'Abstain' is much less intrusive than voting the opposite of your group.

Chapter 8: Finally, in Chapter 8, I have gathered the strands and produced a working conclusion for this paper, to see what we can learn from the analysis, whether some of the research questions could be answered satisfactorily and what I think should be the subject of further research. Two important questions are why I get different results from Hix (2004) when it comes to candidate selection, and what the results have got to say about the electoral accountability of the legislators in the European Parliament.

1.4 A technical note

The paper has been type-set in the Computer Modern font made by Donald E. Knuth using LaTeX. All statistical analyses have been done in R (using the R-Studio GUI), benefiting from various packages written by the R community.

Chapter 2

Legislative Voting

Carey (2009) argues that there are three potential sources of legislative party unity: similar preferences, discipline and agenda control. In addition, Hix, Noury and Roland (2009) argues that division-of-labour and the cost of collecting information can have implications on how cohesive parties are. I will review the arguments in the following sections by first providing a basic model, and then elaborate on this.

2.1 A basic model

Within a legislature, we find party X. X has six members, called X1-6. X1-6 are going to vote. What determines their vote? A simple model would say that the vote is determined by the legislators preferences regarding the vote outcome. In a majority system, if 8/14 legislators want A, then 8/14 will vote for A and A will be enacted. Here the party has only the minimum function of providing the electorate with a common label for a group of candidates. However, if the legislators were randomly assigned to membership in party X, then giving a vote for X in the election would be meaningless – you could not know that the group you were voting for actually voted for policies you would like. Thus, we should believe that parties at least functioned by structuring preferences. After the structuring process, we could know that X1-6 had X-like preferences. Still, in this model, for any given vote, X2 may want something else than X1. X2 and X1 would in that case vote dissimilarly.

Let us then say that party X has a leadership. The leadership have some means to pressure the legislators into voting something they would not have, had it not been for the pressure. The leadership of party X want X^*. X1-4 also want X^*, but X5-6 want another alternative. However, because of pressure, X5-6 choose to follow the rest of the

party and vote for X^*. The party leadership do not bother to apply this pressure in all cases. For instance, in that same case, if they knew that all members in party Y, also a six member party, would vote for X^* (called Y^* in Y) and that the last party Z, a small 2 member party, where against X^*, pressuring X5-6 would not make a difference. The result would be either 10/14 or 12/14, and in both cases X^* would be accepted. Because of the cost (money, time, loyalty, etc.) of applying pressure, party leaderships would rather not pressure when they do not need to.

After a while, X1-6 find they are using way too much time learning about each case they are voting on, only to find that – in general – the other members of the party come to the same conclusion as themselves. Thus, they decide to divide the cases between themselves, and vote whatever the legislator who got the case decides is the best. Now, when a legislator votes, her own preferences are only relevant if she knows she has a different opinion without having been able to investigate the case more closely.

Party X also has another trick up its sleeve. It can actually decide, in cooperation with party Y (or party Z or both), which proposals that are going to be voted over in the first place. Neither party X nor party Y want to put a specific political case on the agenda (in Norway, the best example is EU membership), either because they know it would be splitting the party into factions, or because the electorate would punish them in the next election if that specific case where to come at this time.

If put into a bad spot, seeing that their preferred solution X^* would be punished by the electorate in the next election, party X may decide to vote strategically. This should especially be the case if they know they cannot even win the vote. Say, if Y and Z conspired to get an issue on the agenda where Y and Z want $\overline{X^*}$, and where X^* will be punished by the electorate. From Y and Z's position, this is a perfect issue to have on the agenda. In this case, party X are forced to follow their beliefs X^* knowing they will lose and get punished by the electorate, or vote $\overline{X^*}$. In some cases, it may be advisable to vote $\overline{X^*}$, opening up the possibility of strategic voting on party level.

Individual preferences, party discipline, division-of-labour and agenda-setting combined with the possibility for strategic voting, are thus the main structures in determining what legislators will vote. Parties are motivated by two things: Getting their policies accepted, and winning elections. Individuals may have more complex desire-structures, for instance altruistic goals. However, it is not unreasonable to assume that legislators want to be re-elected, or at least get other positions provided by the party, and follow their policy preferences. The main strategic choices will thus be between individual preferences and party boons, and party/individual preferences and electoral boons. In addition

there will be a cost-benefit calculation in gaining knowledge and in whether to pressure party-representatives and punish rebels.

This very basic model is situated in a simple legislature with simple parties. There are great variations on how parties work, and how they produce representatives and manage them. Furthermore, there may be other groups included into the model. For instance, we could add a president or a government with agenda-setting power or other means to control the vote (like a veto-power). In the European Parliament, the nationally elected parties organise into party-groups. Adding to the equation the sheer amount of national parties represented, the massive amount of legislators, from 27 countries – and thus 27 more or less independent national media that are arguably most interested in national matters during election (Hix and Marsh, 2011), special grants to the party groups, a complex committee system with strong interest group pressures, and a bicameral quasi-parliamentary system with complex procedures, it is easy to appreciate the complexity in this system (Hix and Høyland, 2011, p.56). To disentangle this, I will move through each of the structures in depth, especially noting what is relevant for the European Parliament.

2.2 Preferences

The bottom-line argument when it comes to the relationship between the legislators' individual preferences and party unity, is that if everyone in a group agree to what should be done, then there are no reasons to expect divergent actions by the individuals in that group. In such a case, there is no need for discipline or other arguments for why the legislators follow the party line. We should, however, ask how these preferences are made, and whether the party can be instrumental in forming these preferences. If it can, then there may be differences in how well parties manage to form their legislators' preferences in a way that improves coherence/unity.

Despite being mentioned as perhaps the most important reason for party cohesion, I find little theory about the formation of similar preferences within a political party mentioned in the party cohesion literature. There are at least two main ways in which parties contribute to producing individuals with similar political preferences. First, by providing political alternatives in the public debate. Herein lies a two-way interaction between public attachment to political parties and (re)production of political elites. Second, by being an organisation with active members and deliberate recruitment, socialisation and selection.

A provider of political alternatives. Political parties are providers of political alternatives in the public debate. This is communicated in two ways. Either, through what Leon D. Epstein calls 'the programmatic function' (Epstein, 1980, p.261f), or through political action and statements by political stakeholders bearing a party label. A programmatic party is a party with 'a settled long-range program to which the party is dedicated in definite enough terms to mark it off from rival parties' (Epstein, 1980, p.262). He notes that parties do not provide this function to the same degree. For instance, the Democratic party in the US are to a less degree holding such a political program than the Labour party in Norway. More generally, programmatic parties were associated with mass-parties, arguably because it was essential for such parties to commit to a program to which loyalty could be given (Epstein, 1980, p.261). However the decline of mass-parties, and arguably the decline of programmatic parties, there are still relevant differences between parties in this dimension.

By providing political alternatives in the public debate, political parties are in many ways framing how we view challenges in society and what credible (and possible) solutions to these challenges are. In that way, political parties are narrowing the political landscape. It is within this landscape and around its borders that politicians in modern democracies must work and future politicians are raised. Basically, if you have totally different ideas than what is possible in the given political landscape, you do not become a politician. Rather, you become a philosopher or revolutionary. There is thus a heavy selection for persons able to work within the given environment, and at the same time a drive in persons who want to wield political power to fit into the given political landscape. By having a programmatic function, a party can passively form and select future party elites into the same mould.

A political programme or ideology can make the decision-process for legislators more easy. I.e. a question about more or less income tax is much more difficult if not taken within the framework of parties. Among politicians (and others), there are more persons who just know whether she wants more income tax because of party affiliation, than who actually have sat down and calculated on all possibilities and found out that more or less income tax is the best (given her other preferences). Thus, political parties can improve cohesion by providing political/ideological cues that make decision-making easier (and more conforming to party cleavages).

By being relatively cohesive groups of like-minded legislators, parties produce a preference-map within the legislature. Here, political parties can have more or less overlapping preferences/alternatives, and parties may overlap in some cases and have different alternatives

in other cases. It is probably not unusual for a legislator to find herself in general agreeing with her own party, but disagreeing in particular cases (while agreeing with other parties). These preference-maps can be more or less predisposed for cohesive parties. In systems where individual preferences maps only to either one of the parties, it is less likely with defections than in a system where individual preferences often map across parties.

Perhaps most important when thinking about political preferences within the legislature, is how a political party is placed on this preference-map versa the other parties. If a smaller party is placed between two larger parties, it might be in a position to be pivotal. In such situations, different groups within the party may want to go in different directions – thus splitting up the party. On the other hand, if the small party is positioned on one of the extremes, being pivotal may actually increase cohesion as the party understand that they have a chance to make a difference.

The party organisation While political parties as organisations in general have shrunk in size since the days of the mass-party (Epstein, 1980), their legislative representatives and a label is not the only thing that remains. Even though youth parties have suffered even more member loss than their mother parties (Hooghe and Stolle, 2005), they are perhaps the prime example of how parties engage with their members in order to reproduce (but also reinvent) themselves. In this process, aspiring politicians are being brought together, working for a party from young age. While already undoubtedly having pretty similar preferences, arguably some extra convergence of preferences should be expected to happen to this group in such settings. We could thus postulate that parties recruiting from party youth organisations will get politicians with more similar preferences than parties recruiting from the general adult population. In a interesting study from Flanders in Belgium, Hooghe and Stolle (2005) found that 41% of the city councilors surveyed had started their career in the youth organisation of their party. Also, councilors that had started their career in a youth organisation had received their first mandate on average eight years earlier than those who didn't start their career in a youth organisation. There is good reason to believe that the youth party serves as an important recruitment-channel for political parties.

Those who select candidates for the election must make choices that will affect how similar preferences their representatives in the legislature will be. In general, three traits should be valuable in a candidate for those who select her. First, the candidate should hold many of the same preferences as the selector. Second, the candidate should be good at transforming these preferences into political results. Third, the candidate should be able to win votes in the election. A selector is often faced with the dilemma of having

to prioritise between these traits, as would-be candidates seldom excel at all three traits. Furthermore, there may be more than one selector, and they can have different preferences. In addition to these general structures of candidate selection, parties use specific methods to select their candidates. However, even though these methods may produce more or less similarly minded representatives, the mechanism proposed is that the candidate selection methods improve/reduce party leadership's power over representatives, and thus party discipline (Hix, 2004; Faas, 2003).

2.3 Discipline

In interviews done by Carey (2009, p.26) in the South- and Middle-American legislatures, most legislators reported that most votes are a matter of discipline. Preferences alone cannot account for the level of cohesion we see in modern legislatures. So, why do legislators that are already elected for a certain period of time choose to discipline themselves instead of following their preferences? Jones (2002, p.158ff) cite two important sources to this discipline: a) to protect or improve their political careers, and b) being in the party is often a fundamental part of a legislator's personal identity and where she has important social relations.

Legislators protect their political careers either by assuring re-election or by trying to secure positions linked to the party. In both cases they need to stay within the graces of the party leadership, or at least a part of the leadership. A breach of discipline may be severely punished in several parties. Carey (2009, p.28f) mentions cases where a single breach leads to legislators having to retire from politics, not being able to aspire to other electoral posts or losing strategic positions within the legislature. Further, if the party leadership does not punish breaches, the electorate may. This is especially the case if the electorate perceives the party to be uncohesive and not able to deliver as promised during the election (Hix, Noury and Roland, 2009).

Re-election is an especially interesting case, because it is not entirely up to the party leadership whether a legislator will get re-elected. There are two important steps that any would-be candidate needs to go through before becoming an elected legislator. First, she needs to be selected by the party to run as a candidate. Established parties often have elaborate rules, both formal and informal, for how this selection takes place. The range of methods is quite large, even within countries (Lehmann, 2009; Hazan and Rahat, 2010). And even though the party leadership usually is a key actor in this step, there are several examples of parties that deliberately deny the party leadership any decisive control of the

selection process. I will take a closer view on candidate selection processes in Chapter 3.2.1.

During the election itself, some systems give more detailed control over the vote to the electorate than others. Within PR systems, we have the difference between closed list systems, where the voters only can indicate party preferences, and open systems. In open systems, voters can rearrange candidates on a pre-ranked list or have to rank an un-ranked list, or simply vote for a single candidate. It is thus argued that since the party leadership has less control in such open systems, they also have less control over re-election and thus harder time disciplining the legislators (Hix, 2004; Faas, 2003).

In the European Parliament election in 2009, 12 out of 27 countries had a closed system (Duff, 2011). Only Ireland and Malta opted for something else than a proportional list system with an electoral formula. Instead they use the Single-Transferable Vote system. This is a candidate centered, open system where voters get to rank the candidates. If the person you ranked first either gets enough votes to enter the legislature, or does not have enough votes, the vote counts as a vote for the one ranked as number two, and so forth. The main advantage with the STV system is that fewer votes get wasted. Just as important, voters can rank persons less likely to get elected at first position, because they are not risking that their vote will get wasted.

Scarrow (1997) investigated where the Members of the European Parliament (MEPs) ended up after the next election in a time-series from 1979 to 1989. Her conclusion was that the European Parliament was beginning to attract delegates who served long European careers. However, for 28% of the MEPs investigated, the EP was a political dead end, not leading either to a European, nor a domestic political career (although, other careers may have opened up). We should be careful to draw any conclusions about the current state of affairs on the basis of this research. It is interesting to see, however, that even then, a political career within the European Parliament was something more than $\frac{1}{4}$ of the MEPs realised - showing that a position in the EP is more than just a stepping-stone on the career-path and that re-election is a highly valued alternative for the MEPs.

2.4 Agenda Control

If a party has control over what kind of bills that are submitted to the floor, then it is reasonable to think that it does not want submit bills where the party is split. Thus, we will observe higher cohesion in parties that have control over the agenda-setting than in

parties that do not.

Cox and McCubbins (2005) model how agenda-setting works when opening up the black-box that is the party. Their key assumption is that majority parties delegate to their senior partners the power to set the legislative agenda. These then try to establish a "procedural cartel" that ensures that its members get all/most of the agenda-setting offices and expect their agents to obey "the first commandment of party leadership": To not aid bills that will split the party (Cox and McCubbins, 2005, p.24).

In most parliamentary systems around the world, the government introduces most legislative bills which are mostly accepted by the parliament without large changes (Rasch and Tsebelis, 2011, p.270). This is the case for, among others, Germany, Greece, United Kingdom and Norway. The picture is not entirely clear, though, as legislatures in France, Hungary, Portugal, Spain and Switzerland were found to give legislative bills more pounding in the parliament. In their book, Rasch and Tsebelis (2011) argue, in addition to the importance of the position of the agenda-setter and the possible win-sets of the status-quo, that some actors (often governments) introduce instruments of agenda-setting. These are time constraints, restrictions on the amendment-possibility, sequencing rules (rules governing when different actors are allowed to make amendments), gate-keeping rules (essentially enabling agenda-setters to refuse to make a proposal), vote rules (both winning-rules and vote sequencing) and lastly the degree of exclusive government jurisdiction.

In the European Parliament, it is the Commission that has agenda-setting powers when it comes to legislative initiatives. The Commission needs the support by a majority in the European Parliament before taking office, and a QMV in the European Council to be nominated (Hix and Høyland, 2011). Thus the Commission can be expected to share basic policy preferences with the majority in the EP. Also, since the Commission actually needs to pass legislation through the EP, it is unlikely that they will propose legislation deemed unlikely to pass through the system. This implies that proposals are made that can be seen to fit with a majority-coalition of parties (Hix, Noury and Roland, 2007, p.111). Even though the EU cannot be said to be a typical parliamentary system, the fact that the Commission and the EP is synchronised implies that when it comes to agenda-setting, it is not unreasonable to think about the EU as a bicameral parliamentary system with a government that provides a legislative agenda that is aligned with the wishes of the majority in the parliament and Council.

When a legislative proposal comes to the European Parliament, it is referred to a committee which assigns a rapporteur for the bill. The rapporteur is responsible for drafting the legislative report and leading the work with the given legislation through

the EP. Rapporteurship is assigned by an auction system, where each political group gets a quota of points, and where each group can only bid a maximum of 5 points for each bill. This assures that the largest groups will get most reports, but also that the smaller groups will get some that are important to them (Hix, Noury and Roland, 2007, p.113). If no group bids for a report, it is usually allocated to the committee chair for 'free' (Yordanova, 2011, p.101). The agenda-setting potential here is great, because the report made by the rapporteur contains amendments on the legislation which in theory may change the proposal completely. Also, even though the report needs to get though a vote in the committee with possible amendments from other members, the report that has passed the committee is submitted to the full plenary of the European Parliament (Hix, Noury and Roland, 2007, p.113f).

In addition to being able to respond to proposals from the Commission, committees can also propose non-legislative resolutions and own initiative reports. These are treated essentially the same way as legislative bills when moving through the system. A rapporteur is appointed, and the report is amended and approved, before it is submitted to the floor for plenary voting. It is here suitable to point out that the composition of MEPs in the committees is representative of the EP as a whole (Hix and Høyland, 2011, p.58). Thus, agenda-setting and the possibility of getting proposed amendments into a proposal is dependent on the committee-majority, which in turn is dependent on the political composition in the EP.

When a report is submitted to the full plenary, any political group can propose amendments to the report (Hix, Noury and Roland, 2007, p.114). This is the only place where minority groups have a chance of putting proposals to vote that the majority did not approve of. The plenary votes first on these proposals, and then on the Commission's draft legislative bill as a whole as amended by the parliament. Rasch (2000) argues that in most European legislatures, including the European Parliament, it is customary "to vote for the most far-reaching or extreme alternatives first and to approach more moderate ones gradually" (Rasch, 2000, p.15f). This is done partly because the EP utilizes a successive procedure, where proposals are voted for in succession, and those which are approved are accumulated into the final proposal which is voted for last. Other parliaments use an elimination procedure where different alternatives are put against each other. However, the successive procedure opens up for strategic ordering of the votes, possibly saving a losing bill (i.e. by including concessions to the nay-voters in an amendment), or killing a winning bill (i.e. by managing to form a majority for an amendment that reduces the desirability of the bill as a whole) (Rasch, 2000).

In summary: With the exception of plenary proposals, the check to agenda-setting power by a majority-coalition is pervasive through the whole process from deciding who shall be rapporteur, to which amendments are accepted in the committees, to which full-plenary amendments are accepted in the full plenary (and to how the votes are ordered) and finally to what the final proposals that are put to the vote are.

An important distinction in the European Parliament is between votes that are settled with simple majority and votes settled with absolute majority. The EP uses absolute majority during second reading in the ordinary legislative procedure (OLP), under the consent procedure if it is about admission of new member states, and for amendments of the Council proposal for the annual budget (Hix and Høyland, 2011, p.68f,224). In all other cases, the simple majority rule is used (the famous exception to this rule is under motions of censure against the Commission, where a double majority rule is used: an absolute majority and two-thirds of the votes cast (Hix and Høyland, 2011, p.44)). It follows from this that the large majority of votes in the European Parliament are settled with simple majority.

Empirical evidence from the European Parliament give some support to the idea that majority-parties try to get control over the legislative agenda. Yordanova (2011) finds that the EPP-ED and ALDE got more co-decision (OLP) reports than the rest. Since there is a centre-right majority possibility in the EP, a possible interpretation here is that the EPP-ED prefers ALDE to get reports as opposed to the PSE. It is also interesting that disloyal legislators are given much less reports than loyal ones (Yordanova, 2011, p.114), showing that agenda-setting power also is given as a boon to loyal legislators who work to keep the party group together (or that disloyal legislators are not trusted the job of drafting reports).

2.5 Division-of-labour

A lone legislator in any legislature faces a daunting task. She needs to gather information about each bill that is being submitted to the floor, write amendments and propose own suggestions on a wide range of subjects. In addition, she needs to bargain for positions and agenda-setting powers while maintaining good relations with media and others. Moreover, she loses out on being able to bargain for more concessions by other groups, because the others do not care about that single vote. If all legislators were lone legislators, then complex modern legislatures would not be possible. However, the one thing that is arguably lost when joining a political group, is the possibility to follow your own

preferences without anyone telling you what to do. Hix, Noury and Roland (2009) argues that division-of-labour and sharing of information costs are the main reasons for why MEPs with similar preferences choose to establish political groups and that division-of-labour and sharing of information costs are the reasons for why these groups stay together in a coherent way.

In this framework, agenda-setting can be argued to be one of these competences that are given to certain members of the group. Another important competence is to be a member of a specialised committee. By dividing the work done in the parliament into specialised committees, MEPs can focus their attention on specific subjects and know that others who share your preferences are taking care of the other subjects. Interestingly enough, even though the committees have similar distribution of group members to the full plenary, expertise is not distributed in the same way. This can be a problem, as is argued by a former vice-president of the PES in McElroy (2006, p.13):

> Committees are definitely and regrettably not representative of the Parliament in plenary, they are not microcosms; this results in legislative distortion. The environ-mentally minded from all groups are on the Environment committee, giving it a distinctly green outlook; likewise there are too many farmers on Agriculture. The result of this specialisation and lack of representativeness is that policy is not reflective of the majority view of the Parliament and we frequently have to spend hours in Parliament voting to correct the committee report and proposed legislation. (Personal interview)

Despite this type of inefficiency, separating the workload into specialised committees arguably lets the legislators work more in depth with each proposal. As a side-effect, party unity is bound to increase, because there are fewer persons who are trying to make up their mind about each single proposal. If others are going to follow the legislator on the proposals she is responsible for, then she should try as far as possible to be loyal to what the others tell her. This should be especially interesting because the area she has been delegated is often one she is especially interested in.

Chapter 3

'Electoral Institutions and Legislative Behaviour'

Do macro-level electoral institutions affect microlevel legislative behaviour of individual parliamentarians? If they do, how? This question is raised by Simon Hix in an article in World Politics from 2004 (Hix, 2004). Hix notes that a suitable case for the investigation of this question would be the European Parliament. The reason for this is that we have in the EP a great deal of variation on electoral institutions, both across countries and parties.

Hix approaches this question with what C. Wright Mills would have called his sociological imagination (Mills, 2000). Imagine that you are a Member of the European Parliament (MEP). Hix assumes that once inside, your goals are winning re-election, securing policy and obtaining higher office. You are connected to two different groups that can help you achieve these goals. The two different groups are your national party and the EPG your national party is affiliated with. The downside about these groups is that they sometimes want to support policies that differ from your preferences, and will try to pressure you in order to make you do what they want. They pressure you by threatening to make it harder for you to obtain your other goals, re-election and higher office. More specifically, national parties control re-election and the obtainment of national higher office. Since re-election is a prerequisite for the two other goals, Hix regards this as the primary goal. The party group controls who gets to be rapporteurs and who gets speaking time in the EP – and thus your ability to individually affect policy, as well as the higher positions within the European Parliament.

The interesting point is that within the EP, the MEPs are selected and elected using different rules according to which party and country they come from. These rules, Hix

argues, affect how credibly a party leadership can pose their threats. The single best story to motivate this point comes from the introduction in a recent book written by Reuven Y. Hazan and Gideon Rahat:

> Michael is the head of a workers' union... [H]e learned that the party leadership had decided to allow all party members to partake in the process of selecting its candidates. Michael quickly realized that since his union's members were also party members, they could now take part in the selection process... When the party members' votes were counted, Michael was near the middle of the party list – a "safe position" – and knew that even if his party faired poorly in the general election, he would soon be a member of parliament... [A]fter the election victory, the leadership realized that there were numerous party representatives who refused to follow their directives. Michael, for one, told the party whip that he was selected as a party candidate due to his abilities to mobilize a personal support base, which he could mobilize again next time, and not due to the influence of the party leadership (Hazan and Rahat, 2010, p.1f).

This story, which has 'strong links to events that really took place', ended up with the party having to spilt and a new election being called. It is a testimony for the claim that the method used to select candidates can have a huge impact on party cohesion. The electoral institutions which Hix proposes will have an effect on the national party's ability to control their representatives are the electoral formula, district magnitude and ballot structure and the candidate selection process.

The electoral system: We can define the electoral system as the electoral formula, district magnitude and ballot structure. The electoral formula is all about how the country should translate the votes into representatives. It can be divided into plurality/majority systems and proportional systems. Proportional systems can again be divided into the single transferable vote system and the list systems. For the EP elections, the plurality/majority systems are not used, and only a few countries use the STV system. The far most popular system is the proportional vote list system. Here again, we have different distribution formulas, based on the ideas of highest averages (i.e. D'Hondt) and largest remainders (i.e. Hare).

More interesting for Hix is the interaction between electoral formula and ballot structure. The ballot structure in proportional votes list systems can be open, semi-open or closed. If it is closed, then voters can only vote on whole lists, making this system

party-centered as the party here has a substantial power to decide the ranking of their candidates. In a semi-open system, voters can either vote for the whole list or for individual candidates, while in fully open systems, voters only vote for individual candidates. Hix argues that closed and semi-open systems should be classified as party centered systems, while STV and open list systems should be classified as candidate centered.

District magnitude is the question of how many seats each voting district controls. Hix argues that the party leadership in districts that controls many seats has a more difficult task controlling the candidates than in small districts where it is easier for candidates to run as independents.

Candidate selection: The candidate selection process can be quite complex as we shall see in the next chapter. In his 2004 article, Hix talks about the candidate selection institution in terms of who decides who should be on the party's list ('The Selectorate'). He uses a survey done by Tapio Raunio of MEPs and national party officials that ask how the MEPs were chosen to be candidates (Hix, 2004; Raunio, 2000). The alternatives were 'central party officials', 'party congress', 'regional party organisation' and 'members' (Faas, 2003, p.862). Based on this, he classified countries as either having a centralised or decentralised candidate selection system.

Hix postulates an interaction between candidate selection and the electoral system. The candidate selection rules are more influential in party-centered systems than in candidate-centered systems, because in party-centered systems, the electorate cannot change the party decisions regarding the ranking of the candidates, while the opposite is true in candidate-centered systems. It is important to note that under party-centered systems, both centralised and decentralised candidate selection systems will gain. In the example with Michael, we saw that he was shielded by a closed-list system.

3.1 Operationalisation, model and results.

The model used by Hix to test the theory was a fixed-effects logit model where the two dependent variables were whether the candidate was considered to have defected from the national party line or the EPG line in a voting session or not. Hix defined a defection as when a MEP voted something other than the majority of the national party (or the EPG if we are talking about defection from that group). Ties and parties with less than three representatives were not included. A MEP can either vote 'yes', 'no' or 'abstain'. If the majority abstained, then one voting 'yes' was considered to have defected.

The electoral institutions were accounted for with three variables. *Electoral system*, which is a dummy that asks whether the system is candidate-centered or party-centered. *Candidate Selection* is a dummy where one is that the candidate is selected under centralised rules and zero is that the candidate is selected under decentralised rules. Hix operationalised this on country level and not party level. Thus, he ended up saying that German parties were decentralised, while the French were centralised. The electoral system-dummy and the candidate selection-dummy were then combined into four different dummy variables to account for the theorised interaction effects. *District Magnitude* is a continuous variable measuring how many seats each district have been given.

To account for the conflict level between a representative's two principals, the national party and the EP-group, Hix included two measurements produced from surveys done by the European Parliament Research Group (EPRG). The first, *Left-Right Distance*, measures the distance between a MEP's national party and her EP-group (based on surveys of MEPs) on the left-right dimension. The second, *EU Integration Distance*, does the same for the EU integration dimension.

Seniority (experience in the EP) was argued to be a good indicator of influence. The more influence, the less susceptible the MEP is to pressures from her principals. Thus, more experience may imply a higher probability for defection. High experience may also imply that the MEP does what it needs to survive, thus having a lower probability for defection. Even further, high experience means that the MEP has been in the EP for a long time. Thus there may be some kind of process of socialisation making the MEP agree more with the EP-group than the national party. Lastly, it can also be an indicator for how pleased the national party is with the MEP (if you have been in the EP in many periods for the same party, it is probably because the party likes your work).

He also included several other control variables. These were dummy variables for each of the large national parties and EP-groups. At last he included dummy variables for each vote, to control for the effects of "varying levels of defection and political salience of each vote" (Hix, 2004, p.211).

The result of the regression was that MEPs were less likely to defect from their national party and more likely to defect from their EPG if there was either a party-centered electoral system or a centralised candidate selection (p.216) Thus, Hix concluded that *"electoral institutions shape the relationship between elected politicians and their local and parliamentary principals"* (Hix, 2004, p.219).

3.2 A critique

Hix ventured on a very interesting research route in the 2004 paper. Testing the effects of macro-level institutions on microlevel action by using the abundance of data we have in the roll-call votes and choosing the European Parliament as the case to get variation in the type of institutions is simply brilliant. In addition, Hix produced hypotheses that are testable and falsifiable, which is imperative in scientific research. Regardless, there are two main problems and one interesting methodological question.

The first problem is that the operationalisation of the candidate selection process on country level is not optimal. As I will show below, there are a lot of variation in candidate selection methods between parties within countries. This defect can lead us to read the results in Hix' paper differently. The effect Hix finds and attributes to the electorate institutions may just as well be explained by national culture's differences as to how politicians should behave towards their party (or something else). This is because Hix never really managed to operationalise *the institutions* in a satisfactory way. There is still a job to be done before we can conclude with Hix that "electoral institutions shape the relationship between elected politicians and their local and parliamentary principals". One of these jobs is to operationalise and test candidate selection more properly.

The second problem is that while Hix differentiates between defection *from* the EPG and defection *from* the NP, he does not do so when it comes to the theoretic explanations. In addition, I argue that the main distiction should not be between defection from the EPG and defection from the NP, but between defection *by* a larger group (say a national party from a EPG) versus defection *by* individuals (say defection by a individual MEP against her national party or both her national party and her EPG).

On the methodological side, I am interested in seeing whether the fixed effects logit is the best model to use here. Firstly, I am generally a bit skeptical about running a logistic regression with more than 1000 independent variables. Further, it may be that the average effect across EPGs/other groups is misleading. Most of the proposed effects are mechanisms working within each party/party group. Thus it makes sense to estimate the models within groups and see whether the effects vary between groups. This calls for multi-level modeling.

3.2.1 Selection of candidates

Reuven Y. Hazan and Gideon Rahat has worked much with the candidate selection process, and in the book "Democracy within Parties" they give a comprehensive account

of the important dimensions of candidate selection. They define candidate selection as "the non-standardized and predominantly unregimented particular party mechanisms by which political parties choose their candidates for general elections. The result of this process is the designation of a candidate, or list of candidates, as the candidate(s) of the party" (Hazan and Rahat, 2010, p.4). One interesting observation here is that there is nothing strict, clear and non-fuzzy about the candidate selection process. There are a couple of reasons for this. The first is that political parties are generally regarded as private bodies. The state does not directly interfere with how they go about their business. Second, within these private bodies, there are persons struggling to come into a position of power, or to help other come in a position of power, or to make sure this or that person or group do *not* come in a position of power. There is a Machiavellian side to the candidate selection process.

In their book, Hazan and Rahat identify four dimensions of candidate selection: a) Candidacy – who can be selected as candidate? b) Selectorate – who selects the candidates? c) Decentralisation – where does the selection take place? d) Appointment and Voting – how are the candidates selected? These four dimensions are then analysed in the perspective of intra-party democratisation.

With democratisation, they mean "a widening of participation in both the supply and the selection process - that is, when parties adopt more inclusive candidacy requirements *and* selectorates" (Hazan and Rahat, 2010, p.31). This corresponds to the two first dimensions. Their argument is that these two dimensions have a rather simple effect on intra-party democratisation. The other two dimensions of candidate selection, on the other hand, do not show the same simple interaction with democratisation. While a decentralised system is often perceived as more democratic than a centralised one, the decentralised system may be controlled by local gatekeepers – thus not functioning to widen the participation. Similarly, voting is often perceived as more democratic than appointment. If we analyse this in terms of representation (the idea that the candidates should mirror the different groups in the party and opinions of the members of the party), some types of voting may end up favour the majority groups in the party at the expense of the minority groups. Proportional systems are most likely to ensure representation, while appointment systems can be anything from a good internal debate and weighing between groups to a homogeneous elite just choosing whoever it wants.

Candidacy Requirements

Candidacy rules can come both from national and party level. Typical rules are age-requirements and party membership. Hazan and Rahat propose that we should see these requirements in an exclusive-inclusive dimension here as well. Thus, if there are no requirements, then we have a very inclusive candidacy. On the other extreme, Hazan and Rahat describe a party where only persons that had been a member of the party, trade union, co-operative and insurance association for at least five years; had made annual minimum purchases from the co-op; had been a regular subscriber to the party's newspaper; had sent his (!) children to state rather than Catholic schools; and had his wife and children enrolled in the appropriate women's and youth organisations was eligible as candidates (Hazan and Rahat, 2010, p.20).

The Selectorate

Perhaps the most important among the dimensions of candidate selection, is the selectorate. Hazan and Rahat operate with an inclusive - exclusive dimension. They identify five archetypical groups that can be the selectorate, and they are placed on this dimension. The most inclusive group possible is all voters, like in American primaries. Next is party members, followed by party delegates, party elites and the party leader.

Unfortunately, it is rarely only one group that make the selection. For instance, a party may have a small group identify possible candidates first, before voting in a national assembly. Hazan and Rahat propose that we should identify different complex methods like the assorted, the multi-stage and the weighted method (Hazan and Rahat, 2010, p.36f). Then the researcher must make a careful consideration of many aspects concerning the relative power over the selection process that different groups have, and how exclusive each group really is, before placing the party at hand on the inclusive - exclusive scale. They advise us to use a 24-point scale to pick up the needed nuances, where the archetypes 'leader', 'elites', 'delegates', 'members' and 'voters' serve as guiding points (Hazan and Rahat, 2010, p.48).

Decentralisation

One way to understand the centralisation - decentralisation dimension is on territorial terms: Local, regional and national. However, as Hazan and Rahat states, there are also other groups that are defined non-territorially that can change our understanding of where the selection take place. Instead, they propose that we should think of a centralised

system as when all candidates are selected by the same selectorate, while a decentralised system is when different candidates are selected by different selectorates. An example of a decentralised system would be the one used in CDU in Germany. Here, each Länder-organisation has the power to produce its own list. The central party leadership does of course try to impact the result, but does not have any formal powers over the process (Lehmann, 2009, p.59ff).

Hazan and Rahat make a point out of the centralisation - decentralisation dimension: Scholars often confuse and mix this dimension with the inclusive-exclusive dimension. Even though a centralised selection process often imply a more exclusive process than a decentralised process, it does not need to be. There can be local, elite gatekeepers and there can be national primaries. I am not convinced about the clear divorce between these dimensions. Take for example the case of Michael. Say that he got selected under a national wide intra-party primary. This was then a centralised selection, since the same selectorate – all members – elected all the candidates. However, it also makes sense to say that Michael got selected by a different selectorate than the rest – his being the union members.

Appointment or Voting

The last dimension that Hazan and Rahat brings to our attention, is how the candidates are selected. They see two fundamental different types here, those being selection through appointment, and selection through voting. They propose that we should consider a system a voting system only when two conditions are met: First, the candidate's votes must be the sole determinant of their candidacy. Second, the voting results must be used to justify and legitimise the candidacy (Hazan and Rahat, 2010, p.72).

A problem with analytically separating the mode of selection and the selectorate is that these are not independent of each other. It is simply not possible to have a system where all party members appoint the candidates. Indeed, this is not even found where national assemblies are put to the task. Further, when it comes down to a small group of elites, the difference between appointment and voting becomes very small. However, it makes analytical sense to differentiate between these modes. It gets especially useful when more closely investigating the differences between the systems that use some kind of vote-system.

3.2.2 Selection of candidates before the 2009 election

In order to appreciate the diversity of candidate selection methods in Europe, I will give some examples from the European Parliament report (Lehmann, 2009).

A selection committee appointed by the national board gathered nominations from members and associated organisations for the Moderates in Sweden. The alphabetic list of nominations was sent to the 26 regional boards, which selected five candidates. If a nominee was selected by three or more boards, then she automatically passed the next step, which was a intra-party primary arranged with the help of a postal ballot. Members could choose five nominees, just like the regional boards. Based on these results, the selection committee made a proposal. They were not required to follow the results of the primary. The committee then presented their proposal to the party council, which made the final decision (Lehmann, 2009, p.334).

In D66, members with support from 1 % of the voters could stand for election. D66 involved the party congress, which presented a draft profile for a list of possible candidates. They then performed two referendums. The first to choose the head of list, and then another for the rest (Lehmann, 2009, p.219).

The Conservative Party first had an 'expanded Regional Selection College' deciding whether sitting MEPs who wished to continue should be exempted from the member vote arrangements. That way, the party leadership had implemented a defensive measure for the incumbents against the party members who do the formal selection of candidates through an intra-party primary. The report did not indicate how possible candidates got nominated. Once selected, the candidates (according to ConservativeHome) could only communicate with members via template CVs edited by Conservative Campaign Headquarters. A further complaint from the party members was that ballots where not all candidates had been ranked were disqualified (Lehmann, 2009, p.352ff).

Citizens for European Development of Bulgaria had a meritocratic twist on their selection. First, as candidacy requirements, CEDB/GERB required that nominees had "knowledge of at least two official, non-Bulgarian languages of the EU, tertiary education and at least five years of relevant professional experience [and] knowledge of EU institutions and polices" (Lehmann, 2009, p.18). Submission of nominees was made through municipal and regional party structures, with 50 eventually being passed to the Executive Council. There, a specially appointed sub-committee consisting of seven members interviewed 30 of these and they had to take a written test relating to the work in the EP. Based on the result of the test, 21 candidates got ranked. Finally, Boiko Borisov, the leader of CEDB (and now Prime Minister), intervened regarding the person leading the

list (Lehmann, 2009, p.18f).

UMP in France used a rather exclusive method. A national selection committee composed of the 30 most important party figures had the task of finding nominees. However, because there was going to be elections in each region (*départements*), it was important to get candidates from the respective regions. They thus interviewed political heads in each region and consulted departmental committees. This had been made compulsory since the 2004 election, with a possibility for the departmental committees to call for a members' vote if they were unhappy. The definitive proposal was then put to vote within the National Council (Lehmann, 2009, p.133).

Meanwhile, the British National Party had developed a pure exclusive system, stating that "Local units must submit their choice of candidate for approval to the National Chairman or a person authorised by him to approve candidates" (Lehmann, 2009, p.350).

In Lega Nord, regional lists are proposed by voting committees of the regional party branches and ratified by the national board (where member delegates, party office-holders and heads of the regional branches sit). In addition, the national board decides the heads of list for each constituency (Lehmann, 2009, p.153p).

These examples show that, with the exception of British National Party, most other parties have a rather complex process which involves many people. Sometimes a member vote is simply a guide for party elites, and other times the party elites guide the vote. Even in exclusive systems, central party elites may not do as they want – in most parties, the power is divided in regions and factions. Moreover, all politicians need to think about the electorate and the end result.

3.2.3 Operationalisation of candidate selection

The largest problem I see with Hazan and Rahat's book is that the inclusive - exclusive dimension, especially for the selectorate, when operationalised, gives the impression of being more objective than it really is. What starts simple and relatively objective (just dividing the parties into the archetypes), ends up with the researcher making many important judgments based on a lot of possible considerations. First, the simple groups are not that simple. I.e. the researcher must decide whether a group is a delegate or elites. Second, in complex systems, she must decide the relative importance between groups.

Third, the recipe is not comprehensive. I.e. in multi-stage methods where members and elites are included, the 24-point recipe does not say what to do. I guess that in a 50/50 case, this party would be coded as a '12', like parties using delegates to decide.

Fourth, it may not always be the case that member-systems are more inclusive than

delegate-systems. Katz (2001) argues that while taking control back to the party leadership would be an obvious method of marginalising the party on the ground, there is another strategy the party leadership can employ. If they expand the selectorate to an even wider group of people, the power of party activists would be diluted. Furthermore, "... the less consistently and intensively involved the participant [is] in the candidate selection process, the more he or she will be swayed by name recognition and the more likely he or she is to take cues from the highly visible central leadership" (Katz, 2001, p.291).

Fifth, there is also a statistical problem with the 24-point scale. The problem is that we do not know whether the scale is ranked all the way up (for a scale to be ordinal), nor whether the effect of moving an arbitrary number of points up the scale is the same wherever on the scale you start (for a scale to be interval). A 24 point nominal code is too cumbersome, the same can be said of a 24 point ordinal variable.

Instead, the coarse division used in the survey done by Raunio (2000) is a good start. The EP report on candidate selection methods shows that there are three main solutions in the EU at present:

1. Selection done by party elites.

2. Selection done by a (specially appointed) delegation/party congress.

3. Selection done by party members voting.

As we have already seen, the real world evidence point to numerous different variations of these main types, and the differences may be hugely important for which candidates end up victorious (or get to stand in the first place). Despite of huge variations in these archetypes, and different motivations for the implementation, we should still believe that a party leadership where a small number of party elites (i.e. themselves) decide who shall be the candidates have greater power over re-election than if the party lets a party congress or party members do the job. It may be, however, that the practical difference for the MEPs becomes irrelevant as the party elites find ways to control member/delegate based systems.

The variation on candidacy requirements between the parties in Europe is not that extreme. The most regular requirements are age, gender (often imposed by the national government, i.e. $\frac{1}{3}$ females) and being a paying party member (Lehmann, 2009). Because of the rather small variations, I have opted not to take this into consideration in my statistical models.

One way to control away those member and delegate systems where it is really the party elites that have the power is to find those vote systems that use a proportional voting rule. The reasoning here is that in majority rule systems, the party elites have made a ranked proposal that is voted over, while under proportional rule systems, the vote decides the ordering of the candidates. Thus, party elites have less control over the list under a proportional vote system than under a majority system.

The complete list of codes can be found in the Appendix. I use three different variables, denoting whether the selection is done by a delegation, party members or neither, as well as whether the party uses a proportional vote system.

Chapter 4

Operationalisations of party unity

The most common measurement of party cohesion is the Rice score, developed by Stewart. A Rice in the 1920s (Rice, 1924). It has the simple form

$$Rice = \frac{|Yes - No|}{Yes + No}, \tag{4.1}$$

where Yes is the number of votes for and No is the number of votes against. It can be applied in situations where the voters can be divided into two groups, those who vote 'yes' and those who vote otherwise. The score ranges in the interval from 0 to 1, where 0 is when there is an equal number of votes in each of the two groups, and where 1 is when all votes go to one side. Unfortunately, in all votes where a simple majority is needed in the EP [1], there are three groups that must be taken into account: 'yes', 'no' and 'abstain'. The reason for this is that under simple majority, those who vote 'abstain' do not count for the total. Thus, we must differentiate between the group that vote 'abstain' from both the group that vote 'yes' and the group that vote 'no'. On the other hand, under absolute majority there must be an absolute majority of all representatives in the EP that vote 'yes' to get the vote through. Here, a vote against, an abstain and not being there in effect counts for the same. An alternative to the Rice-score for the simple votes is the Agreement Index

$$AI = \frac{max(Y, N, A) - \frac{1}{2}[(Y + N + A) - max(Y, N, A)]}{Y + N + A}. \tag{4.2}$$

Here, rather than the number of 'yes'-votes versus 'no'-votes, we are interested in the largest group versus the other two. A party will have the same AI if it votes 4 yes, 2 no and 2 abstain, as if it voted 0 yes, 4 no and 4 abstain, or 1 yes, 4 no and 3 abstain. The Agreement Index is in many ways a generalisation of the Rice-index for more than two

[1] Which is true for the majority of the votes.

relevant outcomes. It is nice because while managing multiple outcomes, it also remains simple.

Desposato (2005) shows through simulation that cohesion scores are artificially inflated for small parties (both Rice and AI). The fix he proposes is to calculate the expected cohesion score given that the party only has two members. That way, all scores are artificially inflated. One obvious problem with this is that none of the scores can be interpreted substantially, because they are all artificially inflated. However, it makes it possible to compare scores between groups of different sizes, something that is useful. The function Desposato proposes we should use is

$$E(C_2|C,R) = \frac{RC^2 + R - 2}{2(R-1)}, \tag{4.3}$$

where $E(C_2|C,R)$ is the expected cohesion score in a party with two representatives given the cohesion score C in a party with R number of representatives.

One big problem with cohesion scores is that they aggregate individual vote data to party level. That way we lose many observations and the nuances between defection from NP versus defection from the EPG disappear. One solution to this is to use a measure of vote defection rather than the cohesion score. Vote defection is the individualistic and vote-specific variant of party cohesion. I define it as when a representative votes against its party goals. So, if vote A is clearly in the party's interests and representative X votes something other than A, then X has made a vote defection from its party. Since this definition goes down to the individual level, it is more precise than party cohesion. Because of this, it also needs more precise observations or assumptions. Usually, we do not have more precise observations, so we just add some assumptions.

The assumption which is usually applied is that whatever the largest group in a party voted was in that party's interests. Individuals who voted differently are considered to be the defectors. This assumption could be argued to be more sound if the largest group has a clear majority in the party and if the number of representatives in that party is higher. In a party of three, two voting 'Yes' and one voting 'No', it may be less convincing to apply this assumption than in a party of 99, where 98 vote 'Yes' and one votes 'No'. When dealing with more than two outcomes, it becomes more complicated.

The coding scheme I use is that a defection occurs if a representative is voting differently from the largest group, and if there are more than one largest group (two or more groups with equal number of votes), a defection still only occurs if a representative is voting differently from these groups. The argument for doing so is that we cannot differentiate between equally large groups. In addition, one could argue that the individual

is acting more in line with its party when it follows any of the majority lines, than if it follows a minority line. This is a bit different from Hix (2004) who simply removed the observations where there was a tie.

Table 4.1: Number of defections, crosstabulation

	EPG Defection		
National Party Defection	False	True	Total
False	530 487	21 320	551 807
	93.62%	3.76%	97.39%
Hix	(85.6%)	(10.3%)	(95.9%)
True	3 630	11 183	14 813
	0.64%	1.97%	2.61%
Hix	(1.0%)	(3.1%)	(4.1%)
Total	534 117	32 503	566 620
	94.26%	5.74%	100%

We can define two main types of defection in the EP. The first is defection from national party and the second is defection from the EPG. These two can then be cross-tabulated to yield four different outcomes as seen in Table 4.1.

Three types of defection is possible in the EP. Defection from the majority in your EPG, but not from the majority in the NP; defection from the majority in the EPG and the NP; and defection from the majority in the NP, but not from the majority in the EPG. I argue that out of these, the main difference is between those instances where a majority in a NP decides to defect from their EPG and those where the defection is from the NP or from both the NP and the EPG. Basically, the first case is the only one that is obviously a collective decision, probably supported and enacted by the party leadership. The point here is that there are reasons to believe that such collective actions have different explanations than individual vote defections.

Thus, as for dependent variables, I use a variable that is true if a MEP defected from the EPG, but not from its national party when modeling defection by a national party; while I use a variable that is true if a MEP defected from the national party when modeling defection by individual MEPs.

If we compare Table 4.1 with Figure 1 in Hix (2004), we can see that there are even fewer defections in my data than in the one from 2000. I have included the relevant figures

for convenience. The major change is that national parties seem even more reluctant to defect from their respective EPGs. The proportion of defections where a MEP voted with the national party majority and against her EPG majority has moved from 10.3% to 3.76%. The two other types of defections have been rather stable.

Three important things can be noted. First, this observation fits nicely with the observation done by Hix, Noury and Roland (2007) that the party groups are becoming increasingly cohesive. Second, the figures should improve our belief in that the defection measurement is actually measuring what we want it to, since they fit nicely with what we would expect. Third, it should lead us to believe that these types of defections are driven by different currents (as argued earlier) since only one of the types has changed drastically.

Chapter 5

Hypotheses

In this chapter I bring with me the perspectives from the previous chapters to produce testable hypotheses about structures or situations that may increase/decrease the probability of party defection in the European Parliament. In a number of these hypotheses, more than one of the mechanisms/structures laid out in Chapter 2 can be relevant. I am not in a position to effectively differentiate between these types, which is unfortunate. However, it may be wise to find out whether there is a strong relationship between the probability of party defection and specific structures or situations before venturing into such an endeavour.

I have five groups of testable hypotheses:

First, there are hypotheses that are dealing with situational differences. Under certain conditions, the same party with the same individuals may be thought to be more or less cohesive because of some situational reasons affecting the structures/mechanisms in Chapter 2. The obvious example here is a different preference-structure within a party between votes, i.e a party may agree on reducing taxes, but not to the same degree on the right for gays to adopt children. Unfortunately, I do not have that type of information for parties and individual representatives in the EP. However, there are other situational differences that I can test. These are close/lopsided votes, final/non-final votes and legislative/non-legislative votes.

Second, there are some hypotheses that have been put forth that are about differences in preferences in a general, non-situational way.

Third, the size of a party is argued to be relevant for party cohesion in a number of ways. Interestingly enough, in the European Parliament a party can be said to have two relevant sizes: that in its national parliament and that in the EP itself. Furthermore, the size of the EPG may also play a role.

Fourth, we have the hypotheses regarding institutional/procedural differences in the selection of candidates and election of representatives.

Finally, it is possible to include the Council-EP interplay and ask whether it should matter whether a party is in government (and therefore has representatives in the Council of Ministers) or not.

Within all these groups of hypotheses, two aspects are important to notice. First, the hypotheses are linked to either one or both of the dependent variables. Second, the hypotheses are all linked to the four sources of party unity: preferences, discipline, agenda control and division-of-labour.

5.1 Situational differences

Close Votes: Krehbiel (2000) argues that in order to get results that we could interpret in a more substantive way from party unity measurements, we must somehow test party unity in an quasi-experimental fashion by fixing all but one factor and get variation on the single factor we are interested in. For instance, if the interesting factor is party discipline, then we must make sure to try to fix all other influences, like party preferences. One way to do this is to look at changes within the same group in different situations. Snyder and Groseclose (2000) take up on this and argues that we should see stronger party discipline in close cases than in cases where there is an overwhelming majority for or against (so-called 'lopsided' votes). They find evidence for party discipline having an effect on the voting behaviour of representatives in the House and Senate in the US. Going a step further, Volden and Bergman (2006) develop a formal model for the US system that also can explain the broad changes in partisanship that appear to take place over time in the US. Benefiting from data all the way back to 1877, they find evidence for party cohesion increasing when parties lost seats, gained majority status, diverged from one another in preferences and if cohesion increased in the opposing party. This research strengthens the belief in that parties use disciplinary methods when needed to come in a position to make a difference. There are a couple of arguments we can make to enlighten both why disciplinary methods are viable within a group, but also why they are not used all the time.

One could argue that there is a cost in applying party pressure. If it does not matter whether one applies it or not, then one should not apply it at all: The party spends unnecessary time and money applying pressure, and the representatives may start to feel that the party is micro-managing them. Rather, by letting the representatives loose

on unimportant cases, the party is gaining goodwill for when it really matters. Volden and Bergman (2006, p.77) have a more rational-choice perspective when they argue that "although each member would prefer to vote without external influences on his or her *own* ideal point, party pressure also influences the voting patterns of others, which may lead to more-preferred outcomes". The bottom line here is that the party leadership have larger incentives to increase party pressure in close votes, and party representatives have larger incentives to follow.

Although some may wish that the European Parliament (EP) was like the House of Representatives, it is clearly not. It is not given that close votes will produce the same pattern in the EP because of the multi-party and multidimensional environment. For instance, when there are three parties, and the third party can tip the vote, close votes may lead this party to be torn apart, rather than be consolidated. Another difference is that the EP has both national parties and party-groups. The effect of close-votes should arguably increase the EPG pressure, while it is not evident that national party pressure should increase. On the one side, national parties will gather their forces under close votes, because the vote is more open, with a possibility for smaller parties to be relevant. On the other side, if the national party is being pressured to do something it does not want to, hiding behind "rebellious" MEPs may be an option.

Also, while EPG pressure should rise, the end result may be negated by national parties who act opportunistically under close votes, thus decreasing EPG cohesion. These are the few times where national parties can be relevant actors, which should lead us to think that they might use that opportunity. EPGs may have a harder time holding their affiliated parties together during these votes.

Hypothesis 5.1 *Close votes will not have the strong effect on party cohesion in the EP as in the US Senate.*

Legislative Votes: Since it is the Commission that propose legislative bills to the European Parliament, the EP has less control over agenda-setting in these cases, than under resolutions that origin from the European Parliament itself. In addition, legislative bills will have stronger political impact than resolutions and possible different impact in different countries, making it more likely that national parties will defect from their party group.

Hypothesis 5.2 *National parties will be more likely to defect from their EPGs under legislative votes than under resolutions and own-initiative reports.*

Final votes: The only possibility for minority coalitions to get proposals on vote is by proposing amendments under the full plenary. There is thus a reason to believe that these proposals are more controversial, and more difficult for the large groups to deal with than the proposal already filtered through the Committees.

Hypothesis 5.3 *It is more likely that a legislator or national party will defect under non-final votes than under final votes.*

5.2 Differences in preferences

Hix (2004) proposes that a smaller ideological distance between the national party and the EPG would decrease the probability for different vote instructions from the EPG and the NP to the MEP in the first place - and thus the probability for defection. This is clearly a theory only meant for defection from the EPG. It makes little sense to believe that defections from a national party should be dependent on the ideological difference between the national party and the EPG, because this parameter does not vary within a party. It is difficult to explain differences within a group by appealing to what the members in the group have in common. The only scenario where this is relevant may be for those votes where the majority is not defecting from the EPG because the NP feels forced to vote with the EPG, but a few (brave) souls choose to stay behind and defect. However, the main force behind this theory should be to explain why national parties sometimes vote different than their EPG.

Ideally, we should have a type-based distance measurement, for instance difference in environmental politics between the EPG and the NP, but that kind of data is not readily available. Thus, both I and Hix use the Left-Right and the EU integration dimensions. While Hix uses data from a survey of MEPs done by the EPRG-project, I will use data from the Comparative Manifesto Project (Volkens et al., 2011). The reason is that the CMP has data on more parties. As argued by Lowe et al. (2011), I will apply their measurement using a different scaling method of the CMP-data based on the logarithm of the ratio between number of i.e. left and right sentences. This scaling method is argued to produce a better fit when compared to expert assessments than the Rile-score proposed by Budge et al. (2001).

Hypothesis 5.4 *A shorter ideological distance between the EPG and the national party will decrease the probability of defection from the EPG by the majority of the national party.*

5.3 Party size

It is the parties with most representatives that have most power when deciding what to vote in the EPG. These parties are more likely to form a majority within the EPG when voting than smaller parties, simply because they have more members. Following, they should be less likely to defect from their EPG, because they are more likely to have decided what the majority-vote in their EPG is.

Hypothesis 5.5 *Large parties in the EPG are less likely to defect from their EPG because they are more likely to control what the EPG is voting.*

Parties can also be large in another way. We could argue that large domestic parties are more exposed in national media than small domestic parties. Following, they will defect more easily from their EPG if national sentiments are against the EPG-position, because they are afraid that voting with the EPG will hurt their re-election.

Hypothesis 5.6 *Large domestic parties are more likely to defect from their EPG because they are more exposed, and pay a higher price if voting against national sentiments.*

Because the large EPGs more easily can form majority coalitions than the small, they control the agenda-setting to a larger degree. They should thus be able to a larger degree than the small EPGs, to filter out all agendas that are likely to split the group. Moreover, by being composed of the core parties in Europe, with a long history, large EPGs may have developed more complete/coherent preference structures that are more readily found in potential candidates. In addition, large EPGs have the power to delegate the top positions within the EP (such as chairs in the Committees), as well as being composed of parties that together hold great power in the EU. Thus, it is not unreasonable that the large EPGs are more able to pressure single parties to hold the EP group line.

Hypothesis 5.7 *Large EPGs are less likely to have defections than small EPGs.*

5.4 Electoral institutions

As shown, Simon Hix argues that candidate centered electoral systems and decentralised candidate selection systems will decrease party leadership control over re-election - thus being less able to discipline the vote. While I will use the same operationalisation of party centered electoral systems, I will use two dummy-variables denoting whether the system is

a member system or not and whether the system is a delegate system or not. In addition, I test whether proportional vote systems are more effective in reducing party leadership control, and thus in increasing the probability of defection from national party, than any of these other archetypes. A proportional vote system is however not independent of these archetypes, as they are only found under delegation and member systems. Thus, they cannot be implemented in the same models.

Hypothesis 5.8 *Candidate centered systems decrease party power over representatives.*

Hypothesis 5.9 *Party leadership has less control in member and delegate systems than in elite systems.*

Hypothesis 5.10 *Party leadership has less control in proportional vote systems.*

District magnitude is the question of how many seats each voting district control. Hix (2004, p.198) argues that the party leadership has a more difficult task controlling the candidates in districts with many seats than districts with few seats, because it is easier for candidates to run as independents in districts with many seats.

Hypothesis 5.11 *A high district magnitude reduce party control over individual representatives, because they can more easily run as independents.*

5.5 EP – Council interplay

It can be argued that MEPs from national parties in governments that are going to vote after a Council compromise (that is, in all budget cases and in legislative absolute majority cases), are being pressured by their parties to vote for the compromise. Thus, they are less likely to defect in the cases where the EP-group are in favour of this compromise and vice versa (Hix and Lord, 1996; Hix, 1997; Faas, 2003). How this turns out in the models is not that clear. To test it in a sufficient manner, we would have to single out those votes that is about agreeing to the Common Position - data I do not have at present. Given that most votes are about agreeing to a large degree with the Common Position and that most EPGs are friendly to this goal, then we should see less defections from the EPG from national parties in government. I will include a variable for whether a party is in government or not in my models, but I do not link this to any hypothesis as I am not in any position to test it.

Chapter 6

Method and data

For the testing of the various theories and hypotheses, I made two methodological choices. In order to get an intuitive grasp of the data, and to show how the data is distributed, I carefully investigate it graphically. Next, I run different types of regression, from a simple logit model, via a so-called 'fixed-effects' model, into a logit-model with clustered standard errors and finally a fully fledged multi-level model. There are some reasons to do so. First, Hix (2004) uses a fixed-effects model, so it is interesting to compare this type of model with a simple model and a clustered standard errors model to see whether they produce different results. Another reason is that the fixed effects model includes more than 1000 dummy-variables, which in itself is dubious. In addition, by including them, the model will not estimate with the situational variables. Thus, using a model with clustered standard errors functions both as a robustness test for the fixed effects model, and as a way to test the situational variables taking into account clustering. Lastly, the data is severely clustered. We have MEPs, political parties, EPGs, countries and votes. The fixed effects model is only removing the between group variation. It may be more interesting to estimate this variation and take into account the between-group variation when estimating our independent variables. This must be done in a multi-level model.

In this chapter, I will first go through my explanatory variables. Second, I will perform a bivariate graphical analysis. In this analysis, I will only note patterns, without formally connecting it to the theory, as I will do so when analysing the results of the regression models. Third, the regression models will be presented.

Table 6.1: Electoral systems for the 2009 EP election.

Countries	# MEPs	Constituencies	Preferential voting	Distribution of seats	Threshold
Austria	17	1	Yes	PR (D'Hondt)	4 %
Belgium	22	3	Yes	PR (D'Hondt)	No
Bulgaria	17	1	Yes	PR (Hare-Niemeyer)	No
Cyprus	6	1	No	PR (D'Hondt / Droop)	No
Czech Republic	22	1	Yes	PR (D'Hondt)	5 %
Denmark	13	1	Yes	PR (D'Hondt)	No
Estonia	6	1	No	PR (D'Hondt)	No
Finland	13	1	Yes	PR (D'Hondt)	No
France	72	8	No	PR (D'Hondt)	5 %
Germany	99	1 (16*)	No	PR (Sainte-Laguë)	5 %
Greece	22	1	No	PR (Variant of Hare)	3 %
Hungary	22	1	No	PR (D'Hondt)	5 %
Ireland	12	4	Yes	STV	No
Italy	72	5	Yes	PR (Hare)	No
Latvia	8	1	No	PR (Sainte-Laguë)	No
Lithuania	12	1	Yes	PR (Hare-Niemeyer)	5 %
Luxembourg	6	1	Yes	PR (D'Hondt / Hagenbach-Bischoff)	No
Malta	5	1	Yes	STV	No
Netherlands	25	1	Yes	PR (D'Hondt)	No
Poland	50	13	No	PR (D'Hondt / Hare-Niemeyer)	5 %
Portugal	22	1	No	PR (D'Hondt)	No
Romania	33	1	No	PR (D'Hondt)	5 %
Slovakia	13	1	Yes	PR (Droop)	5 %
Slovenia	7	1	Yes	PR (D'Hondt)	4 %
Spain	50	1	No	PR (D'Hondt)	No
Sweden	18	1	Yes	PR (Sainte-Laguë)	4%
United Kingdom	72	12	Great Britain: No	Great Britain: PR (D'Hondt)	No
			Northern Ireland: Yes	Northern Ireland: STV	

* Parties can choose to post lists for each Länder. Only CDU did.

Sources: (Duff, 2011; EP, 2012; Lehmann, 2009).

6.1 Explanatory variables - Operationalisation and sources

There is a *Close Vote* when the difference between the number of 'Yes'-votes and 'No'-votes is 10% or less than of the sum of the number of 'Yes' and 'No'-votes. This cut-point has been arbitrarily set. The point here is that when the result was between 40-60% in favour, less was needed to be done in order to flip the vote, than when the result was lopsided. The operationalisation is made after the fact, but it probably captures the relevant votes since the theory is about small groups (like national parties) suddenly becoming relevant. Small groups do not manage to change the outcome dramatically,

taking the end result below 40% or above 60%.

This operationalisation implies that it is only valid under simple majority votes, because for absolute majority votes, 'Abstain' counts as 'No' and should not simply be disregarded, as in the above operationalisation. All models in Table 7.1 and 7.2 in addition to the multi-level model in Table 7.3 using Close Vote as independent variable are estimated using just simple majority votes[1].

The European Parliament are voting both on parts of a proposal and on whole proposals. A *Final Vote* is a vote on a whole proposal. I use the data as coded by Hix et al. (2012).

Delegation Size is coded by counting the number of MEPs in each national party. Since all 736 MEPs[2] are in the dataset, this data should be correct. In order to make the regression models easier to interpret, the variable is centered when used in regression models.

% Votes in National Election have been gathered from the website maintained by Nordsieck (2012). I use the election results that was relevant for Autumn 2009. For Greece, which had an election in the mid of this period, I have chosen to use the 2009 results. The difference between the 2007 and the 2009 result is negligible for my purposes, as the election did not change the composition drastically. Germany also had an election the 27.September 2009. I use the results for after this election. This only has relevance for the 56 first votes in the data set, so I believe this not to be a problem. *In Government* simply denotes the parties that were in government at the time of voting. Regrettably, elections made at a later stage are not included. The relevant countries are Belgium (13 June 2010), Czech Republic (28-29 May 2010), Hungary (11 & 25 April 2010), Latvia (2 October 2010), Netherlands (9 June 2010), Slovakia (12 June 2010), Sweden (19 September 2010) and United Kingdom (6 May 2010), making the changes in these elections only relevant for a small portion of the votes.

Δ *Left-Right* and Δ *EU Integration* is the difference between the mean score on the *logrile* and *logeu* in an EPG and the national party affiliated with this EPG. The variables are provided by Lowe et al. (2011). The latest measurement is used. For some parties, this means that the measurement was taken as long back as in 2002. It is not likely that a party has wandered very far on the left-right spectrum in that time.

A country has a *Party Centered* election system if it uses the STV system or opens

[1]While not strictly necessary in all models, I made this decision in order to make the models easier to compare. The far majority of the votes are done by simple majority.

[2]In December 2009, the EP got enlarged by 18 members. These are not included in the data-set. This is a possible source of error for all results in my paper.

up for a preferential vote. Farrell and Scully (2010) operate with a more nuanced system (dividing mixed systems (the voter can both vote on party and candidate) into those where the candidate vote matter and those where it doesn't), but I have opted for a simple solution. *District Magnitude* is the number of representatives that shall be elected from a country divided by the number of constituencies. Germany is a special case here, since the CDU chose to run for election in each Länder, rather than having a country-wide list. The data for these variables have been gathered from a EP Report (Duff, 2011). In the report, Denmark is noted as not having a preferential vote. I have changed this as it is incorrect (Lehmann, 2009). The figures can be seen in Table 6.1.

6.2 Bivariate graphical analysis

In the Appendix, there are bivariate plots on all independent variables versus a count-version of the two dependent variables. In addition, I have plotted the largest parties (15 or more representatives) against the dependent variables. These plots, in addition to being interesting in themselves, are important as a verification tool towards the regression estimates.

I use two types of plots. The first is a scatter-plot, with colours separating MEPs in different EPGs. The second is a beanplot (Kampstra, 2008). The beanplot is an alternative to a boxplot, but better because it does not hide the data[3]. Instead, there is a small vertical line for each observation, a density curve and a red line for the median observation. If more than one MEP has the same number of defections, the small vertical lines are put together into a longer line, eventually reaching a set maximum. However, the density-curve will depict the data in that sense more properly.

Defection within EP groups

Starting with Figure 8.4, we can see the two types of defections plotted against EPGs. We can see that the NI and EFD-groups have much higher defection-counts than the other groups when it comes to defections by MEPs being in the group who voted what the majority in their national party voted (henceforth 'defections by the national party'). On the other hand, when it comes to defections by MEPs from their national party, these groups do not differ from the other EPGs. At least for the NI-group, this is just as predicted. The parties within this group do not have a goal to be united when voting,

[3]Which is a problem if the data does not conform to a uniform distribution or if there is very little data.

and this is reflected in the votes. Individual defections from the NP is another matter. Here, the NPs in the NI group do not differ from other NPs.

While the average defection-count in defections by the national party is higher than the average defection-count against the national party, there are much more zero-counts in defection by the national party than from the national party. Indeed, no MEPs have zero defections from their national party.

The general trend seems to be in accordance with Hyp 5.7. Large EPGs tend to have fewer defections by national parties than small EPGs. The only exception here is Greens/EFA, which has an extremely low defection count.

Lastly, and perhaps most importantly for the following analysis, we see that there are very little data in the GUE-NGL, NI and EFD groups. This is mainly due to the fact that the national parties in these groups are often small, with less than the 3 members necessary to be able to compute defection scores and that the groups themselves are small. Due to this fact, we should be very careful in interpreting results coming from these EPGs. Also, since many of the mechanisms proposed presuppose that the groups have a chance of obtaining political power, they may not be relevant at all in these groups. In the regression models, I have chosen to remove these groups.

Defection and size of party

Figure 8.5 shows the two size variables plotted against the two dependent variables. There is clearly a difference between the plots using defection by the national party and the plots using defection from the NP. In the upper two plots using defection by the national party, there actually seems to be a trend, while in the lower two, using defection from the national party, the defection-counts seem to be spread in a more uniform fashion. This is just as hypothesized in Hyp 5.5 and 5.6.

Large delegations seem to defect less from their EPG than small delegations. This has probably to do with these parties actually controlling to a large degree what their EPG is voting. We see, however, that there are quite a number of large parties associated with the EPP, all of them having relatively few defections. Thus, this core group of parties must somehow manage to agree, and no single party are controlling the vote. It may be argued that there is a curved, rather than linear effect of delegation size if we look at each EPG for themselves. This seems at least true for the two large EPGs and ECR. Again, this may be because the small parties really do not have any chance of affecting the vote outcome. They are better served by being loyal and getting important positions within the EPG and the EP.

If we remove the EFD and the NI from the picture, there seem to be a general trend that small domestic parties seldom defect from their EPG. In large domestic parties, the picture is more mixed. The proposed theory say that these parties may defect when national sentiments are pressuring them to vote differently than their EPG. Small parties do not get this pressure because they are too small to make a difference anyway. The variation in small parties might be because many of the large national parties also have large delegations in the EP.

Defection and preferences

In Figure 8.6 we see the same tendencies as in Figure 8.5: A possible trend in defections by national party from the EPG but not in defections from the NP. This makes sense. We should really not see a trend in the difference between the ideology in a national party and the mean ideology in the EPG when it comes to individual defections from the NP, because there is no variation between these differences within a national party.

For the two plots on the top in Figure 8.6, it makes sense to disregard EFD and NI. There are many reasons why these should defect from their EPG, so we should not be surprised that they are off the curve in this figure (or any other).

The coding also produces what probably are meaningless results for the smaller EPGs, as the mean gets more affected by the individual parties as the number of parties goes down. This point is probably aptly represented by the Greens/EFA. In parties where the mean ideology probably has a more substantial meaning, like in EPP and S&D, the trend is pretty clear: Increased ideological distance leads to more defections. ECR on EU distance is probably the most beautiful example of this, with a more or less perfect linear trend-line between the three parties represented.

Electoral institutions

Moving to the electoral institution variables in Figure 8.7 and 8.8, the trends are more difficult to find. One reason for this is that they are mostly dichotomous, and have for that reason less possibility for variance. The proposed theory predicts more defections by the national party from the EPG in candidate selection systems where the party leadership has control (non-member/delegate systems) and less defections by MEPs from the national party in such systems. The only variable that seems to show this trend is in delegate systems for defections by the national party from the EPG. Member System and Proportional Vote show the opposite trend in these types of defections. For defection by MEPs from the NP, none of these variables seem to make a difference. Moving on to the

election variables, no trend is apparent in these either. Generally, the bivariate analysis does not bode well for Hyp. 5.8, 5.9, 5.10 and 5.11.

6.3 Regression models

The logistic model

In a logistic regression, we model the probability that $y = 1$,

$$Pr(y_i = 1) = logit^{-1}(X_i\beta), \tag{6.1}$$

out of the two possibilities 1 and 0, under the assumption that the outcomes are independent given these probabilities (Gelman and Hill, 2007, p.79). In order to be a sound model, all the normal assumptions (validity, additivity and linearity, independence of errors, equal variance of errors and normality of errors (Gelman and Hill, 2007, p.45f)) in a regression model must be made. Still, we can only address covariance, and not causality. In this paper, I do not assume specific causal relations, only tendencies in which I propose possible explaining causal factors. It is these tendencies we first need to find out whether have some support in the data.

The $logit^{-1}$-function maps the linear model $X_i\beta$, which can run from $-\infty$ to ∞, to the probability, which must be between 0 and 1. The inverse-logit function is not a linear function, implying that the predicted probability will not change equal amounts with a change from i.e. 1-2 in X and a change from 2-3. Rather, changes in X have the largest effect on the probability at 50% probability of $y = 1$, while it deteriorates as the probability moves towards 0 or 1. This is important to notice, especially in my case, since defection is a rare event, and thus unlikely. Seemingly large effects will produce small changes in the predicted probability in such cases. Gelman and Hill (2007) recommends calculating predicted probabilities and interpret these instead of the effects in logit or odds-ratios because of this.

Standard errors in a logistic regression represent estimation uncertainty. Logistic regression uses a maximum likelihood estimator. The likelihood is defined as "the probability of the data given the parameters and inputs" (Gelman and Hill, 2007, p.388). Thus, the ML estimator tries to find the parameters that is a most likely estimation given the data. In such a model, there are many possibilities that are almost as likely, and they are (hopefully, and given a well specified model) randomly scattered around the most likely estimate. Based on the distribution of these, the model will calculate the deviation of likely estimates from the most likely estimate. Since the Central Limit Theorem states

that the sum of many small random events will be a random variable approximating the normal distribution, and (since properly specified) all that is left in the model are random events, we argue that we will find estimates ± 2 standard-errors away from the most likely estimate in 95% of the samples we collect. This is of course not what we want, since there are no more samples to be taken. We have the full population of data. What we can say, however, is that the estimates are consistent with the data to the degree specified by the standard errors. Since we have not sampled the data randomly from a population, we cannot really infer from the data to other possible data (like voting data for the year after) - other than if we assume that the other data will be similar to this data. One solution would be to run a Bayesian statistical model, which would tell us what we should believe about this phenomenon given our prior belief (which could be specified as a uniform distribution or something similar that is vulnerable to external influence.)

The simple logit model assumes that the residuals, which is the difference between the expected value and the observed value[4] (Gelman and Hill, 2007, p.97) are independent of each other. There are many reasons to think that this is not true in the simple model. The most important reason is that each individual vote action has been made in the context of a vote over a specific proposal. All votes over a specific proposal cannot be said to be made independent of each other, and none of the variables in the simple model can control for this fact. Both the fixed effects model and the GEE models try to answer this problem, but in different ways.

A fixed effects logit model is simply a logit model where we include dummy-variables for different groups in the estimation. The effects for each group thus indicates how the intercept is changed between the base-group (which is arbitrarily chosen) and that group.

The GEE model

The Generalised Estimation Equation (GEE) model is in a way a generalisation of the generalised linear model (GLM). Where the GLM assumes that the observations are uncorrelated (when controlled for the independent variables), the GEE includes a matrix of 'working correlations' across a given cluster for a given set of vectors of observations when calculating the variance of these vectors of observations (Zorn, 2001). That way, the variance within each vector of observations is controlled with a correlation structure within the cluster, and given that this structure is specified correctly, the variance estimate will be consistent. However, even if it is not specified correctly, the GEE-model also includes a robust estimate of the variance-covariance matrix analogous to the Huber-White sandwich

[4] $residual_i = y_i - E(y_i|X_i) = y_i - logit^{-1}(X_i\beta)$

estimator. Thus, the variance estimate will be consistent even under misspecification of the correlation matrix. By using the robust estimate, we are however putting the cart in front of the horse, caring more about the proper specification of the model than on getting a correct estimation of the process that generated the data (Freedman, 2006). The difference between the simple logit model and the GEE model is however the difference between an incorrect model fitted with improper specification and an incorrect (at least if the correlation-structure is incorrectly specified) model fitted with proper specification.

The two models are estimated with a so-called 'exchangeable' correlation structure. Here, I assume that the correlations are the same for all observations within a cluster (Lam, 2007). The other reasonable alternative I had was to use a 'unstructured' working correlation, where no constraints are placed on the correlations. The results are relatively similar, but predictions were a bit better when using the exchangeable correlation structure. As a last note about GEEs, goodness-of-fit statistics are not available due to how they are estimated (Zorn, 2001, 476f). They are therefore not included.

The multi-level logistic model

The multi-level logit model can be generally written as

$$Pr(y_i = 1) = logit^{-1}(X_i\beta + \alpha_{j[i]}). \tag{6.2}$$

The difference between this model and the basic logit model is $\alpha_{j[i]}$. This is the group-level predictor, which in itself is a regression of the group coefficients

$$\alpha_j \sim N(U_j\gamma, \sigma_\alpha^2). \tag{6.3}$$

Written in common words, the group-level predictor is estimated as a normal distributed function of a matrix U_j group-level predictors and a vector γ coefficients, with a σ_α^2 standard deviation of the unexplained group-level errors (Gelman and Hill, 2007).

One trait that is important to notice when it comes to multi-level models versus fixed effects models (or what Andrew Gelman and Jennifer Hill calls 'no pooling' models), is that the estimate of a given group (say 'a vote') in a multi-level model can be approximated "as a weighted average of the mean of the observations in the [group] and the mean over all [groups]..." (Gelman and Hill, 2007, p.253). This has the possible advantage of constraining estimates from groups with few data towards the mean of all data – instead of accepting an estimation based on very few data, which is done in a no pooling/fixed effects model. The multi-level model is called a 'partial pooling' model because it is placed in the space between a basic logit model ('complete pooling'), which only takes into account the

mean effect across groups, and a fixed effects model ('no pooling') which, for each group, only takes into account that group. In a group with little data, the multi-level model uses the mean effect to a large degree, while in a group with a lot of data, the multi-level model uses the group-specific effect to a large degree. This might come in handy in the EP-data, since group-sizes differ and there might be some votes with few participants. Especially when it comes to votes with few participants, we would be ill advised to take heed of a possible extreme effect to the same degree as with votes with many participants.

Chapter 7

Results

In this chapter, I present the results and make some notes about how the models compare to each other and the general predictive power of the models. Lastly, I make a short investigation of how often the MEPs use the option to vote abstain when they defect, and also if there are some differences in the use of abstain between the EPGs.

The general result is that the EPGs and national parties are in very cohesive, and none of the structural or situational variables change this radically. The strongest predictor is differences in ideology between national party and EPG on defection by national party from their EPG. There are some differences in impact of situational variables across groups. National parties affiliated with ALDE have a higher probability of defection than national parties from the other party groups. This is sensible, as ALDE is in the middle of the policy space.

Table 7.1 shows four different logit models, with the simple logits used in the two first, and the GEE model is used in the two last. The difference between these two first models is that I have included all votes as dummy-variables in the second model. In addition, I had to remove the situational variables in order to be able to fit the model because by including single votes as independent variables, and having variables describing groups of votes, estimation was done on data with no variation on the dependent variable (I.e. in a specific vote, what is the effect of that vote being a final or non-final vote on the dependent variable?). This leads to meaningless results. All models are only estimated on votes using the simple majority rule, because I use the close vote variable. Furthermore, I had to reduce the number of observations, because the estimation got too demanding for the computer. I have drawn 100 000 random observations from the data, out of which 76 614 had complete information. These observations are used in the fixed effects model and the two GEE models.

Table 7.1: Logit models

		Def. by NP		Def. from NP
	Simple Logit	**Fixed Effects**	**GEE Logit**	**GEE Logit**
(Intercept)	-3.380 ***	-3.405 ***	-3.340 ***	-2.676 ***
	(0.103)	(0.733)	(0.140)	(0.129)
Situational variables:				
Close Vote	0.103 ***		0.165	0.191 *
	(0.024)		(0.115)	(0.102)
Final Vote	-0.584 ***		-0.354 **	-0.523 ***
	(0.023)		(0.129)	(0.103)
Legislative Vote	0.194 ***		-0.343 **	-0.090
	(0.019)		(0.134)	(0.093)
Preferences:				
Δ Left-Right	0.492 ***	0.507 ***	0.453 ***	-0.130
	(0.040)	(0.106)	(0.121)	(0.120)
Δ EU Integration	0.101 ***	0.085 *	0.078	-0.071
	(0.014)	(0.036)	(0.049)	(0.046)
Size:				
Delegation Size	-0.053 ***	-0.056 ***	-0.050 ***	0.019 ***
	(0.001)	(0.003)	(0.005)	(0.004)
% National Election Votes	0.033 ***	0.035 ***	0.032 ***	-0.018 ***
	(0.001)	(0.003)	(0.004)	(0.004)
In Government	-0.058 *	-0.037	-0.031	0.139 *
	(0.023)	(0.061)	(0.068)	(0.073)
Electoral Institutions:				
Member System	0.210 ***	0.195 ***	0.187 *	-0.410 ***
	(0.022)	(0.058)	(0.083)	(0.078)
Delegate System	-0.133 ***	-0.113	-0.111	-0.207 **
	(0.030)	(0.076)	(0.105)	(0.087)
Party Centered Election	0.320 ***	0.378 ***	0.339 ***	-0.171 **
	(0.021)	(0.055)	(0.077)	(0.064)
District Magnitude	-0.003 ***	-0.003 *	-0.002	-0.000
	(0.000)	(0.001)	(0.002)	(0.001)
EP groups (ALDE is baseline):				
ECR	-0.100 **	0.054	0.037	-0.042
	(0.036)	(0.095)	(0.135)	(0.120)
Greens/EFA	-2.328 ***	-2.636 ***	-2.463 ***	-1.139 ***
	(0.100)	(0.282)	(0.329)	(0.160)
PPE	-0.569 ***	-0.530 ***	-0.488 ***	-0.776 ***
	(0.032)	(0.083)	(0.113)	(0.110)
S&D	-0.822 ***	-0.781 ***	-0.708 ***	-0.886 ***
	(0.029)	(0.076)	(0.095)	(0.097)
N	453800	76614	76614	76614
-2LL	-63483.94 (df=17)	-8275.379 (df=1236)		

***: $p(H_0) < 0.001$, **: $p(H_0) < 0.01$, *: $p(H_0) < 0.05$

The table shows four different logit models, the first three with the same dependent variable. I have divided the independent variables into the groups as indicated by the hypotheses. All models are estimated with data from simple majority votes only. Because of higher demands on the computer, the three last models are estimated with less data. The fixed effects model has in addition 1221 dummy variables, one for each vote, but they are for obvious reasons not printed.

Table 7.2: Multi-level models

		Def. by NP		Def. from NP		
		Base Model	**Full Model**	**Base Model**	**Full Model**	**Using Prop. Vote**
Random effects:	Name	Variance	Variance	Variance	Variance	Variance
Votes	(Intercept)	4.018	4.098	0.953	0.879	0.879
		(2.004)	(2.024)	(0.976)	(0.938)	(0.938)
EPGs	(Intercept)	1.064	0.856	0.216	0.259	0.329
		(1.031)	(0.925)	(0.465)	(0.509)	(0.573)
Fixed effects:		Estimate	Estimate	Estimate	Estimate	Estimate
	(Intercept)	-4.947 ***	-4.663 ***	-4.091 ***	-3.766 ***	-3.764 ***
		(0.466)	(0.425)	(0.210)	(0.232)	(0.260)
	Close Vote		0.585 **		0.251 **	0.251 **
			(0.190)		(0.090)	(0.090)
	Final Vote		-0.713 ***		-0.553 ***	-0.553 ***
			(0.151)		(0.072)	(0.072)
	Legislative Vote		-0.294 *		-0.193 **	-0.194 **
			(0.147)		(0.070)	(0.070)
	Δ Left-Right		0.533 ***			
			(0.039)			
	Δ EU Integration		0.049 ***			
			(0.014)			
	Delegation Size		0.029 ***		0.011 ***	0.008 ***
			(0.006)		(0.001)	(0.001)
	Delegation Size^2		-0.0025 ***			
			(0.0002)			
	% National Election Votes		0.029 ***		-0.020 ***	-0.014 ***
			(0.001)		(0.001)	(0.001)
	In Government		-0.085 ***		0.145 ***	0.029
			(0.022)		(0.024)	(0.022)
	Member System				-0.479 ***	
					(0.027)	
	Delegate System				-0.221 ***	
					(0.031)	
	CS:Proportional Vote					-0.464 ***
						(0.029)
	Party Centered Election				-0.113 ***	-0.186 ***
					(0.023)	(0.023)
	District Magnitude				-0.0003	-0.001 *
					(0.0004)	(0.000)
	AIC	118862	108364	107142	106486	106555
	N	494389	453800	494389	494389	494389
	Votes	1223	1223	1223	1223	1223
	EPGs	5	5	5	5	5

***: $p(H_0) < 0.001$, **: $p(H_0) < 0.01$, *: $p(H_0) < 0.05$

All five multi-level models are estimated allowing the intercept to vary between EPGs and votes. In the second model, I have included Delegate Size^2, because it fits the data better that way. Because of this, Delegate Size is not centered in this model. I only estimate the models using those variables that I deem relevant for each dependent variable.

Table 7.3: Varying slopes: Situational variables

	(Intercept)	Close Vote	(Intercept)	Legislative Vote	(Intercept)	Final Vote
ALDE/ADLE	-4.153 ***	1.276 ***	-3.680 ***	-0.787 ***	-3.637 ***	-1.011 ***
	(0.065)	(0.174)	(0.064)	(0.081)	(0.066)	(0.131)
ECR	-3.824 ***	-0.601 ***	-3.512 ***	-0.847 ***	-3.897 ***	0.591 ***
	(0.065)	(0.192)	(0.065)	(0.083)	(0.068)	(0.128)
Greens/EFA	-6.690 ***	-3.099 ***	-6.667 ***	0.020	-6.594 ***	-0.007
	(0.112)	(0.891)	(0.120)	(0.136)	(0.120)	(0.216)
PPE	-4.706 ***	0.898 ***	-4.515 ***	-0.214 **	-4.274 ***	-1.301 ***
	(0.064)	(0.172)	(0.063)	(0.077)	(0.065)	(0.129)
S&D	-4.680 ***	0.746 ***	-4.386 ***	-0.696 ***	-4.246 ***	-1.671 ***
	(0.064)	(0.174)	(0.064)	(0.079)	(0.066)	(0.133)

***: $p(H_0) < 0.001$, **: $p(H_0) < 0.01$, *: $p(H_0) < 0.05$

The table depicts three different multi-level models, where the relevant variable have been allowed to vary between EPGs. The intercept are allowed to vary between EPGs and votes. Then, I have extracted the estimated coefficients and their standard errors for each EPG, which is what you see in the table. The models use all votes, except the one with Close Vote as independent variable, which only uses simple majority votes.

In Table 7.2, we can see the multi-level models. I have made one base and one full model for each of the dependent variables, only allowing the intercept to vary between groups. In addition, I have included a model where I use Proportional Vote instead of Member/Delegate system. In all models, estimations are done within two different types of groups. First, within each of the five (out of eight) EPGs and second, within all votes[1]. In these models, I have only kept those variables which I deemed to be interesting explaining factors for defection by national party and defection from national party respectively.

Further, in Table 7.3 I have made three different varying slope and intercept multi-level models. In each model, I have taken the base model with defection by national party as basis, and just included one of the three situational variables. These models can find out whether close votes, final votes and legislative votes have different impact on party cohesion for each group. The table shows only the estimated coefficients for both intercept and the relevant situational variable, instead of the estimated variation.

Situational variables: While the multi-level model estimates the variation between groups, it is also possible to extract estimated coefficients for each group. Especially interesting here is to plot the estimated coefficients for the votes, because as Table 7.2 shows, there is huge variation between votes on defection by national party. In Figure 7.1, I have plotted the estimated intercept for each vote, with the estimated standard errors, for the first base model. These must be interpreted as the expected probability for the average EPG in each vote. For around 400 of these, the expected probability is

[1]But only those using the simple majority rule, as I will be using the Close Votes variable.

close to zero. Examples of such votes are VoteID 1181 (search in votewatch.eu) "Anti-Counterfeiting Trade Agreement (ACTA)", VoteID 1298 "Control by Member States of the Commission's exercise of implementing powers" as well as all votes under the headline "A new strategy for Afghanistan" (VoteID 1311-1338). The two first are examples of votes where all MEPs were in agreement, mainly because the cases were about control of the Commission, something all MEPs could agree on. The last example is probably an example of a case where members within a party-group have aligned preferences.

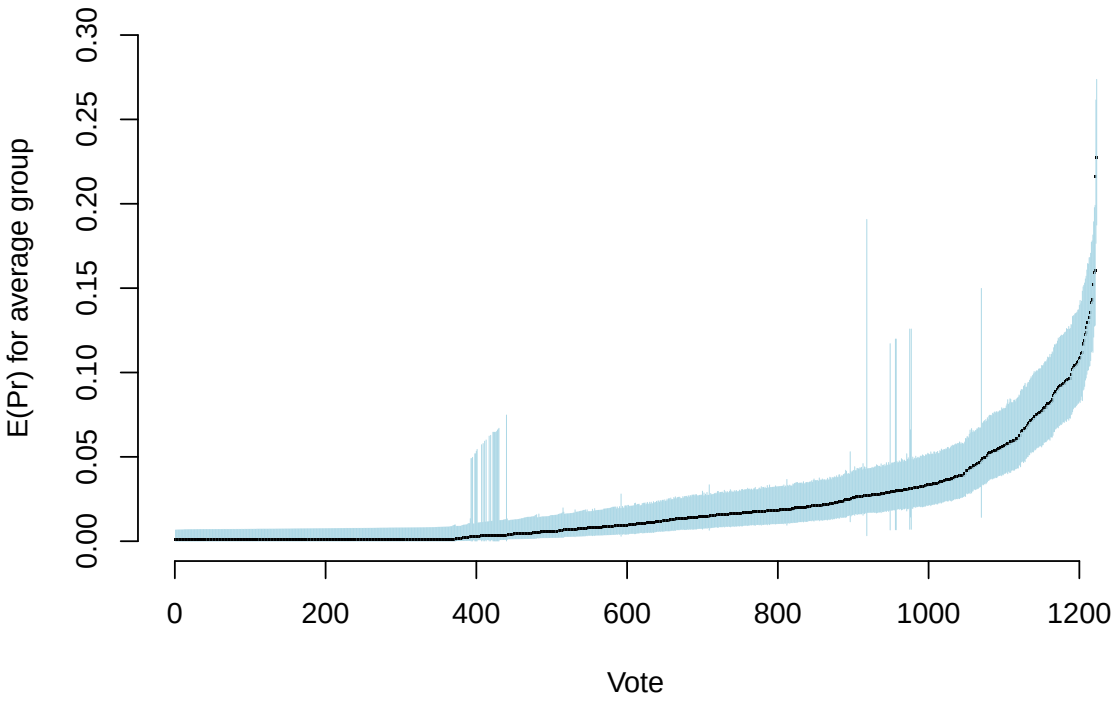
This is in contrast to the two cases with highest estimated probability: VoteID 33-48 "Motions for resolutions - Crisis in the dairy farming sector" and VoteID 1039-1065 "Improvements in the safety and health at work of pregnant workers and workers who have recently given birth or are breastfeeding". In the last example, there was a divide between EPP-members in the Southern-European states and the others.

The large differences in expected probability between votes, and this small peek into specific cases, should lead us to believe that the best explanations for vote-defections are the case at hand and how preferences are divided between national parties in party-groups regarding this case. This, however, should not lead us to not be interested in how these cases are affected by different types of structures and more general situations. It also shows that most votes are uncontroversial and not prone for defection by national party.

In Hyp 5.1, I hypothesized that close votes should not have a strong effect on party cohesion. In the US Senate, it was shown by Volden and Bergman (2006) that party cohesion increased under close votes. If this is the case in the EP, we should see negative estimates. We can see that the effect is estimated to be positive in all the models in Tables 7.1 and 7.2. However, in the GEE Logit model with defection by national party the estimate has a high standard error, and is not significantly different from 0. It is thus not entirely clear that a close vote has an effect on party cohesion.

When looking more closely at the estimates in Table 7.3, however, we can see that the average (positive) estimate from the full model is composed of negative estimates from ECR and Greens/EFA and larger positive estimates for ALDE, EPP and S&D. The estimates make sense if we place the party-groups in a policy space. Greens/EFA and ECR are on opposite sides, in each extreme, while ALDE is positioned in the middle, between the two large groups. This estimation predicts that when the vote is close, the parties placed in the center of the policy space may decide to wander from the position of the majority in their group, to the other position. ALDE is by far most likely to find itself in such a position, as we can see better in Figure 7.2. The red line shows the predicted probabilities in each group as we move from non-close votes to close votes (with

Figure 7.1: E(Pr) in each vote, base model (only EPG defection).



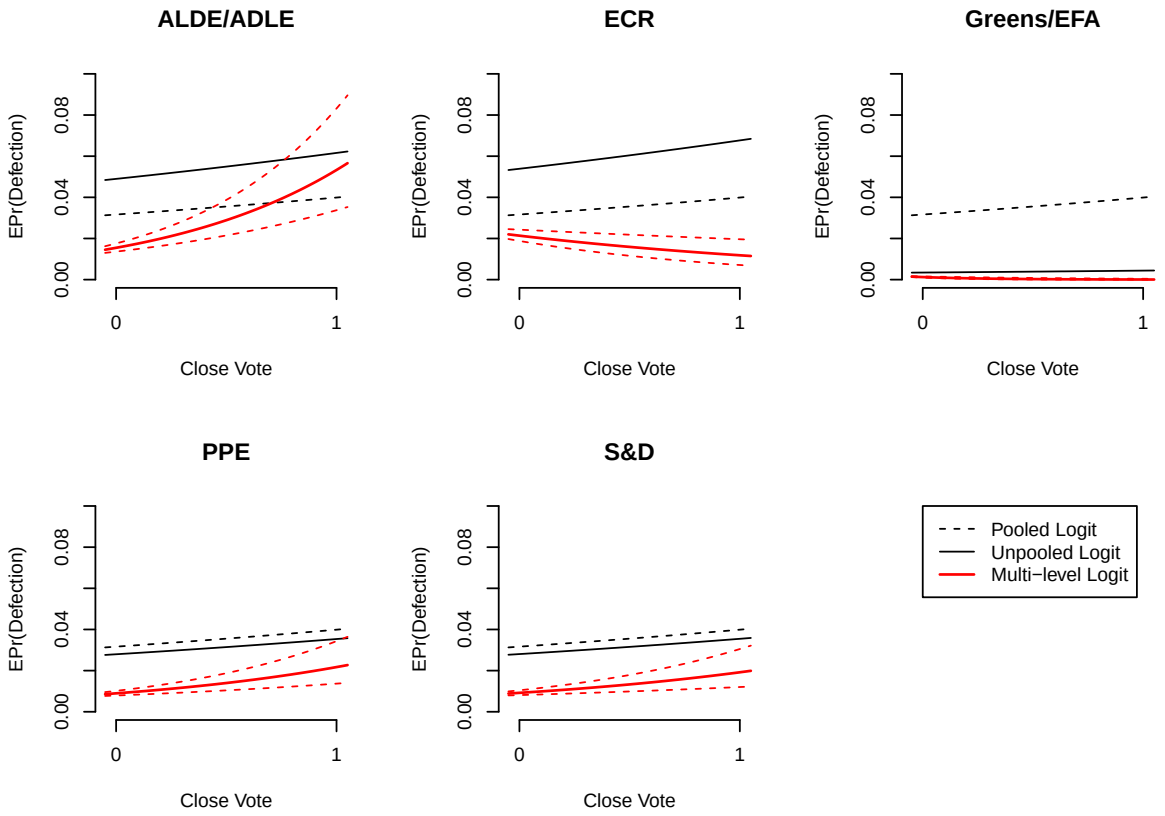
The figure shows the estimated probability of defection, and the 95% Confidence Interval, for all simple majority votes, given the base multi-level model with defection by national party as dependent variable.

the dotted red lines showing the 95% Confidence Interval). The two other lines are simple logit (pooled) and fixed effects (unpooled) logit versions of this model. As we can see, in close votes, ALDE has about double risk (1.5-4) of defection by national parties compared to lopsided votes. None of the other party-groups seem to be affected in the same way.

The main point here, that close votes have a different effect on the EPGs than they have on the two parties in the US, should thus be relatively clear. Moreover, the disrupting effect of close votes seem to increase for the parties in the middle of the political spectrum, while it decreases for those at the extremes.

Legislative votes were hypothesized to increase the probability of defection, because the Commission has agenda-setting power, making it more probable that some unwanted theme would hit the floor. The only model who show this tendency is the simple logit model, while all other models are estimating a negative and significant effect. In this case, we should not believe the simple logit model. The varying slopes model shows that all parties get an expected probability of defection at about 1%, reduced from about 3% for the ECR, and a bit less for the rest. The Greens/EFA do not get an effect, as they already

Figure 7.2: Close votes: E(Pr)



These plots are a graphical presentation of the first model in Table 7.3 (red line). These estimates are plotted with the 95% Confidence Interval (dotted red lines). In addition, this model can be compared with a simple logit ('pooled') and a fixed effects logit ('unpooled') model using only Close Vote as independent variable.

have a very low estimated probability of defection (0.1%). Thus, we can conclude that the hypothesis is wrong, or at least negated by other influences that reduce the probability of defection. Since legislative votes are often more important, it may be that the EPGs demand increased discipline from the national parties.

Parties that did not get their proposals through in the Committee, may put their proposals to vote in the full plenary. These kinds of votes are coded as non-final votes in the data, because the final vote is always a vote over the full proposal, as amended by the EP (which contains the rapporteur proposal and all accepted full plenary proposals[2]). I hypothesized that MEPs or national parties would be more likely to defect under non-final votes, as these probably are more controversial, not being in the rapporteur proposal to begin with.

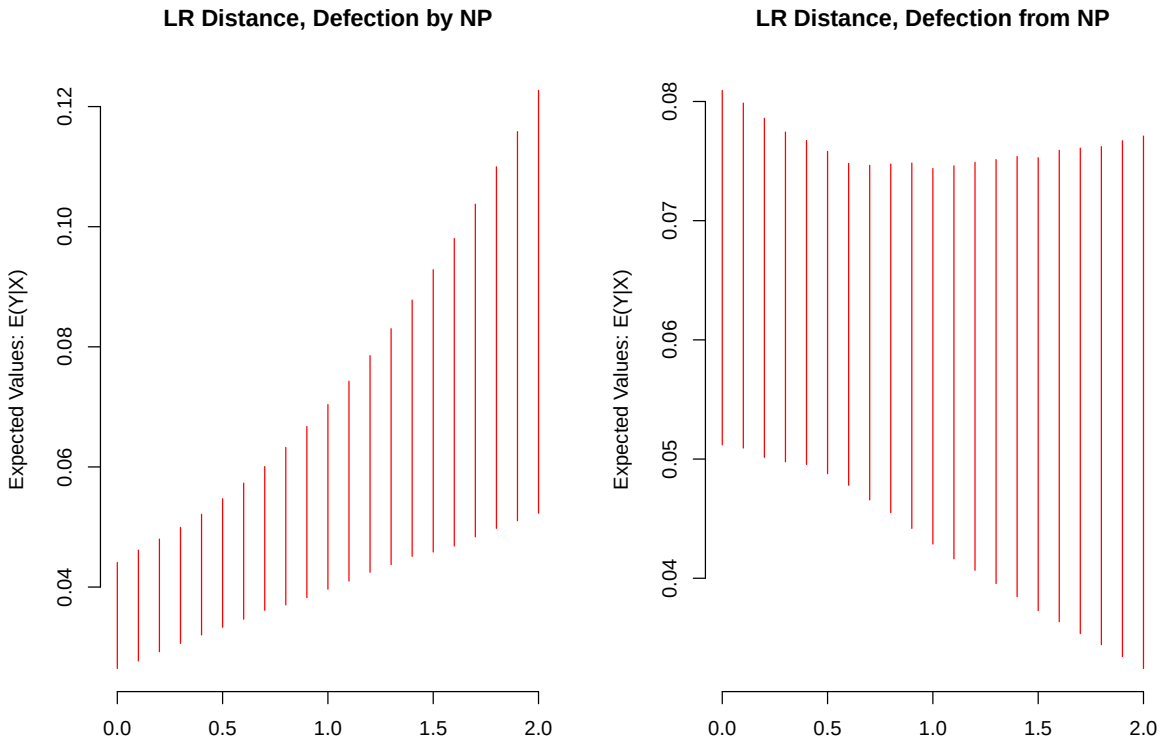
In the tables, we can see that final votes are estimated to reduce the probability of defection, just as hypothesized. This is true for both defection by NP and from NP.

[2]If any at all.

When looking more closely, an interesting pattern emerges. The large EPGs, EPP, S&D and ALDE/ADLE have an estimated reduced probability for defection, while ECR is estimated to be more probable to get defections. Thus, it is evident that the effect first and foremost hits the large parties that are controlling the Committee proposals.

Preferences: I hypothesized that differences in preferences between the NP and the EPG should matter for defection by the NP, but not from the NP. Basically, the more different ideology a national party has from the average in their group, the more likely they will be to defect because the probability of having different preferences in a case will increase. All models show this pattern to a strong degree when it comes to the Left-Right ideological dimension. In addition, the effect is not significant when using defection from the national party as dependent variable. EU Integration, however, do not always have a significant effect, nor is the effect very strong. I thus conclude that Hyp 5.4 is only likely for Δ Left-Right.

Figure 7.3: Expected values for Δ Left-Right, GEE models.



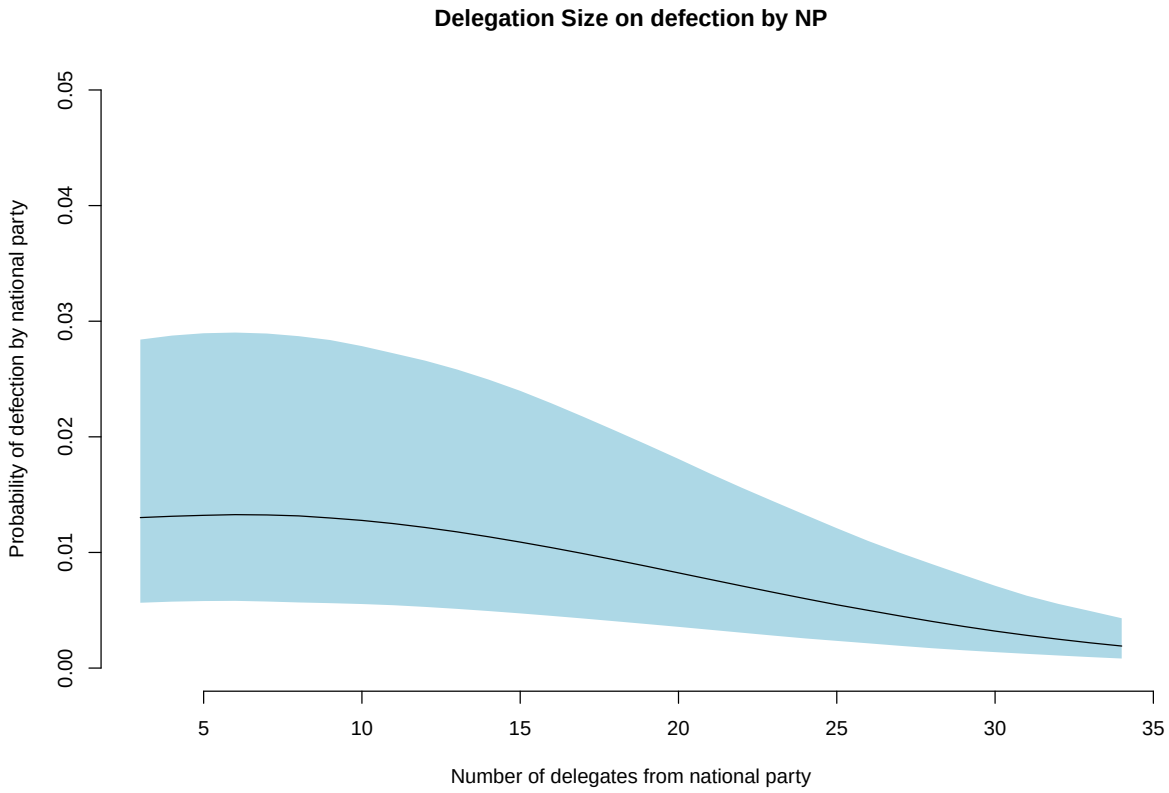
The graphs show simulated expected values given the two GEE models using X=0 for all independent variables except the intercept and Δ Left-Right. Delegation Size, % Votes in National Election and District Magnitude are centered in the models, with a mean of 16, 29% and 25.5 respectively.

In Figure 7.3, I have plotted simulated predictions for the two GEE-models, only

allowing Δ Left-Right to change. This shows two points. First, that Δ LR only has an effect on defection by national party. Second, that this effect is relatively large, however with higher uncertainty as the distance is increased. This, of course, is because there are fewer data to estimate on as distance increases (which also can be seen in Fig 8.6).

Size: Large national party delegations will control many votes in a EPG, thus being more likely to control what the majority is voting than small national party delegations. This is partly because of how we define party defection, but as we saw in the graphical analysis, for EPP, there are several large parties with few defections that need to cooperate. Thus, an effect of this variable is more than just a product of the operationalization. The effect of Delegation size is significant and negative in all models with defection by national party. This variable goes between 3 and 34, so while the marginal effect of moving 1 up or down the x-axis is low, the difference between large and small delegations is estimated to be quite large. In addition, when estimating for the multi-level model, the curve-linearity we saw in the graphical analysis seem to be find support.

Figure 7.4: Expected values for delegation size, multi-level model.



The graph shows simulated expected values given the multi-level model with defection by national party as dependent variable, using X=0 for all independent variables except the intercept, Delegation Size and Δ LR. Δ LR is set to 0.5 in order to get a higher expected probability. % Votes in National Election and District Magnitude are centered in the models, with a mean of 29% and 25.5 respectively.

The sum product of this effect is plotted in Figure 7.4. The plot is made by simulating coefficients drawn from a multivariate normal distribution, using the estimates and their covariance as guides. These coefficients are then combined with chosen values of X and run through the inverse logit function in order to get the predicted probabilities of defection. The blue area denotes the where 95% of these predictions where found to be, while the curve shows the median prediction for each value of Delegation Size.

The graph shows that there is huge variation within parties of the same size, but that this variation is expected to be lower in the large parties, than in the small. Further, size only seems to reduce the probability of defection as party delegation size gets above 20 (or so).

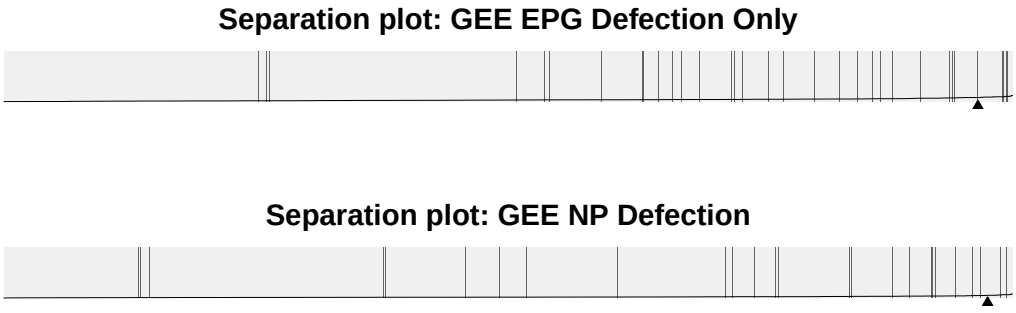
The other way for a national party to be large is to be large in the national parliament. I hypothesized that parties that are large in the national setting should be more likely to defect from their EPG than others (all else equal), because they pay a higher price if voting against national sentiments. This theory also seems to find support in the models, as all estimates are positive and significant. However, this effect is less than the effect of delegation size.

Both these variables have opposite and smaller effects when it comes to defection from the national party. I do not propose any hypotheses for this pattern, other than noting that the pattern is different than from defection by national party. This supports the notion that the operationalisation of party defection I have proposed makes sense.

Electoral Institutions: Electoral institutions were hypothesized to affect the ability of national party leaderships to make credible threats towards their MEPs, by making them more/less able to control the reelection of the MEPs. Decentralised/inclusive systems like member and delegate systems where hypothesized to increase the probability of defection from the national party, while a party centered election and low district magnitude where hypothesized to decrease the probability of defection from the national party. Vice versa, Hix (2004) argues that these institutions should also work in defections by the national party from the EPG by increasing the leadership's power to make their MEPs forgo rewards given by their EPG principals and instead listen to the leadership. Thus, the probability of defection should be opposite when it comes to defection by the national party.

The only variable which is following the hypothesized pattern is Party Centered Election. Member and Delegate system is consistently significant but with opposite effect from the one hypothesized in all cases except for Delegate systems in defections by national parties. Proportional Vote, which should be an even more precise operationalisation of

Figure 7.5: Separation plots



The separation plots show the predictive capabilities of the two GEE-models. A model with perfect prediction would place all vertical black lines on the right of the arrow. Tendencies for clustering toward the right show improved prediction capabilities from a random prediction.

when party leadership has little control over candidate selection, also gets a relatively high negative estimate. It performs consistently as the Member System variable does. District Magnitude has a very precise estimation of 0 effect.

With these estimates, it is difficult to believe that the electoral institutions generally play a significant role when MEPs are deciding what to vote.

Predictive power and model comparison

The standard errors are difficult to interpret between the simple logit model and the other two in Table 7.1. The simple model has been run using six times more observations due to the increased complexity in the other models demanded more than my 8 Gb of RAM could give them. If we compare a simple logit model using the same observations as in the fixed effects model (not shown), the simple logit model gets a -2 Log Likelihood value of -10667.22 on 17 degrees of freedom. Knowing that a random variable tends to increase the -2LL with 2, the change from -10667.22 to -8275.379 by adding 1223 independent variables is not an improvement of the model fit. The coefficients and standard errors of the explaining variables between these models using the same observations are very similar, not changing anything. Thus, in the fixed model, adding the votes as dummy variables is not a significant improvement of the model.

In Figure 7.5 I have made a separation plot for each of the two GEE-models. These show how well the models predict the outcomes in the data. Each dark vertical line is a defection. The arrow points to the expected number of defections as predicted by the model, while the black horizontal line represents the expected probability of a defection as given by the model (Greenhill, Ward and Sacks, 2011). Given a perfect fit, all the dark lines would be on the right side of the arrow. For both models we can see that the

dark lines are spread out all over the plot, with a slight clustering towards the far right. Also, never in this model does the predicted probability of defection become higher than the predicted probability of non-defection. Thus, the model never predicts, even in the most extreme circumstances in our model, that a MEP is more likely to defect than not in a vote. The spread out dark lines implies that the model is not managing to predict the outcomes very well. However, there do seem to be a slight clustering towards the right, and the predicted probability is increasing slightly in both models, so the models are not completely irrelevant either. This should be about what to expect given the data we are using to predict the defections. Except from the situational variables, none of the variables change between the votes. Since MEPs have chosen to be in a party and the parties have chosen to be in a party group, we should believe that the members tend to vote together. Thus, variables capturing the situation as it all the time, should not lead to models predicting defections more often than not. Viewed in that way, the separation plots show that the variables pick up trends in the data which are relevant for prediction, but not in any way do the variables determine defection.

7.1 Abstention

Table 7.4: Defection when voting abstain

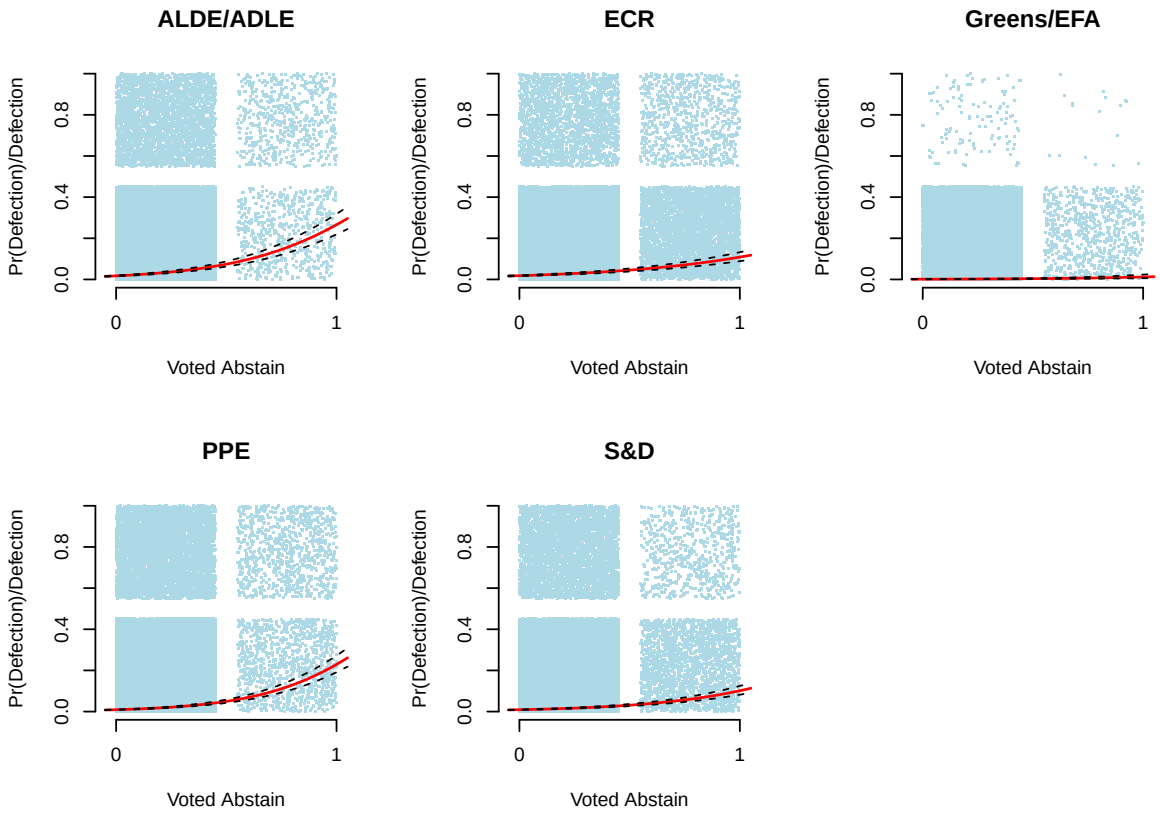
		Voted Abstain	
		False	True
Defection from national party	False	98.1%	1.9%
	True	75.7%	24.3%
Defection by national party from EPG	False	98.2%	1.8%
	True	81.3%	18.7%

One interesting question to pose is whether the option to vote 'abstain' is one used primarily when a MEP or national party does not want to vote what its group want it to. At least in simple votes, voting abstain is a nice option, because such a vote does not count for the total. Under absolute majority, it depends on whether the group wants to vote 'yes' or 'no', since 'abstain' in these cases count as a 'no'. In Table 7.4, we can see that 24.3% and 18.7% of the defections are made by voting 'abstain' in the two types of defections respectively. Furthermore, it is more usual to vote 'abstain' when defecting from national party than from the EPG only. This use of abstention shows that many of the parties/individuals who defect do so in the most non-intrusive way possible. On the

other hand, the majority of defections are defections where the defector choose to vote for the opposite alternative.

In Figure 7.6, I have made a varying intercept and slope multi-level model with Voted Abstain as independent variable, and defection from EPG only as dependent variable. As we can see, the use of 'abstain' when defecting is varying across groups – being most usual in ALDE and EPP. It is interesting that S&D has such low usage of abstention, as it is reasonable to assume that the larger and more powerful EPGs should be able to demand the use of abstention more often than the smaller ones.

Figure 7.6: Voted abstain: E(Pr)



The figures show the expected probability for defection by national party for each EPG given that a MEP voted abstain or not. The red line shows the estimate, while the dotted black lines show the 95% Confidence Interval. These models where estimated on all votes and not only the simple majority votes, as for the rest of the models.

Chapter 8

Conclusions

I have argued that it makes more sense to investigate the European Parliament in terms of party defection rather than party cohesion. Both the national parties and the EPGs are in general extremely cohesive – making a defection a rare event. Furthermore, it seems plausible to believe that national parties collectively choose to defect from the EPG in certain votes, and that the reasons for why they choose to do so (or rather, why they choose *not* to do so when they would have liked to) are part different and part easier to predict than when individual MEPs choose to defect from their national party.

I think the data is clear in showing that EPP and S&D have structural benefits when it comes to party unity compared to the other parties. Their ability to provide the credible policy alternatives and to control the agenda-setting make them stable performers in most situations. ALDE, on the other hand, is in the middle of the policy space, with a possibility of being pivotal. This is slightly increasing the probability of defection by the national parties from the majority line under close votes. ECR seems to have a different explanation for their defections, probably stemming from actual disagreement between national parties at slightly more instances than in the other party groups, which should not be surprising given that they are newly founded.

However, the main tendency in all groups, including the Non-Inscrits, is that there is an extremely high cohesion. The main explanation for this is that as a individual MEP, you are insignificant – more so than in any other legislature in the world. In order to be successful as a legislator in any thinkable way (getting your policy preferences realised or improving your career), you need to work with and for a group.

8.1 Accountability

The extremely high cohesion in the EPGs is worrying so long as the electorate is voting for national parties. In order to be minimally accountable, the EP representatives should be elected under their EPG banner. For instance, under the electoral campaign in Denmark, neither the Social Liberals or the Liberals were actively promoting their mutual affiliation with the ALDE/ADLE group (Lehmann, 2009, p.53). The Liberals were in government at the time, while the Social Liberals were in opposition, so to the general public, they would look similar (both liberal), but different – making it probable that many voters where trying to decide between these parties. Looking at the general cohesion in the EPGs, not informing the public about this affiliation reduces accountability because it probably does not matter whether you vote for the Social Liberals or the Liberals in the EP election.

One way to think of the European Parliament is to view the national parties as individual legislators in national parliaments. The cohesiveness of the national parties is near perfect, and defections from them seem to come at approximately the same rate for all legislators, making the end product look like a random process. Thus, voting for a party because you like a rebel legislator in that party is rather unproductive. In order to be able to be a relevant actor at all, that rebel will need to vote similarly to what the others in her party are voting. One example here is Anna Hedh from the Social Democrats in Sweden. In 2004, as a EU-skeptic candidate, she was placed far down on the list by her party (even on the reverse side of the ballot), but thanks to preferential votes, she got her place in the EP (Lehmann, 2009, p.339). She did manage to stand out, for instance being only one of three Social Democrats who voted against the Lisbon treaty (votewatch.eu, 2008). However, she has voted with her national party 95% of the votes she has participated in (votewatch.eu, 2012). Even though voting for Anna Hedh is a more EU-skeptic alternative than voting for other Social Democrats, you are not first and foremost voting for a EU-skeptic alternative when voting for Anna Hedh – you are voting for the S&D-group in the EP, and they are not EU-skeptic. There are of course other ways individual legislators can influence EU policies and the direction of EU, but this is one of the most important ways they do, and this should be communicated to the electorate.

8.2 What about the electoral institutions?

Hix (2004) and Faas (2003) find that electoral institutions, including the candidate selec-

tion methods, have an effect on the probability of party defection. In this paper, using data from 2009-2010, while finding effects that are consistent with the data, they are opposite from what hypothesized when it comes to the candidate selection variables. This need some explanation.

The best explanation is that most of the parties using member or delegate candidate selection systems are well established parties from central EU member states that has both defined EU policies and found their place in a EPG. Parties like the UK Conservative Party and CDU in Germany. Another important group is most of the green parties, all in a party group with exceptionally high cohesion. Here, the placement on the left of the S&D group is probably not conducive for defection. Thus, even though there may be a negative effect from a inclusive candidate selection system, it can be negated by other effects.

It may even be reasonable to argue that party leaderships only will give delegates or members the control when party unity is high enough for this to be viable. If this was the case, then cohesive parties would be correlated with inclusive member systems – which is the case.

Another possibility is that the effect found by Hix is spurious. For instance, the use of a measurement on country-level may be invalid. It may have picked up on something else of interest, like a political culture. However, Faas (2003) is using the same data, only on party level, and he gets the same results. More probably, the (at that time) recent inclusion of Sweden and Austria, which both have many parties using a decentralised selection, in addition to the UK Conservatives belonging to the EPP-ED at the time, may have produced a picture where the decentralised/exclusive parties were less cohesive than the centralised/inclusive ones. Thus, there are some ways of looking at the data in which the explanation given may not be true.

There are many reasons why the effect of leadership control of reelection should *not* be of any importance for party unity, and even more so as the EPG cohesion goes up. First, individuals are insignificant in the EP, thus the chance of getting candidates that does not want follow the party is very low – and then also the need to threaten legislators. Second, because each EPG consists of many national parties, a rebel or two in a couple of parties really does not matter in the big picture. They may even be positive additions, able to fetch more voters for the party, because they appeal to a different group of voters. For instance, a Swedish Social Democrat party without Anna Hedh would probably fetch fewer votes than one with. Since the risk of a rebel is spread out over 82 different electorates and ten-folds of parties, the parties in a group should not be too worried about a couple

of rebels.

On the methodological side, the conclusions from Hix and this data shows the problem of making generalised claims based on statistical results that is not drawn from a random population. It is not possible to make inferences to other data, other than by assuming that other data is no different than this data – and to argue why this should be the case. There are no statistical reasons to infer to other data.

8.3 Shortcomings

The major shortcoming in this analysis is that I treat all MEPs who do not come to vote as missing data. In Figure 8.10 I have plotted these instances for all national parties in the data. As we can see, the median MEP is not going to vote in about 10-20% of the votes possible. This is quite a high number, and would probably affect the results if we were to code these as either a defection or not. The problem here, of course, is how we should treat these instances. On the one hand, it might be that not showing up to vote is a convenient way to defect. On the other, there are many reasons for why a busy politician in Brussels cannot show up to vote, that has nothing to do with the intention of vote defection. In order to implement these non-voting occurrences into the analysis of vote defection, we need more information about the explanations for non-voting, which is a question too large for this paper. The conclusions should thus be read with the reservation that non-voting does not change the results.

The second shortcoming is that I do not manage to sufficiently differentiate between the structures/mechanisms that produce party unity and deter/provoke party defection. Smarter tests or qualitative design are needed in order to find out how especially party size affects party defection (we now should be more convinced that it do).

As of yet, the explaining variables are too large for the phenomena they are trying to explain. In order to get a fuller grasp of party defection, we will need to delve into specific votes or groups of votes and make statistical models there. This paper has demonstrated that the use of multi-level modeling is a fruitful approach, however the models I have produced fall short of explaining the phenomena they were set to explain.

Finally, the data is very clustered. Using maximum likelihood to estimate on models with more than a couple of these rapidly increases the complexity of the problem and demands more of the data. While I ideally would have liked to estimate the data within all clusters – MEPs, national parties, proposals, votes, EPGs, countries, regions, etc. – this simply has not been possible. There are thus much that can be done to improve our

understanding of legislative voting.

Bibliography

Bowler, Shaun, David M. Farrell and Richard S. Katz. 1999. *Party Discipline and Parliamentary Government.* Ohio State University Press.

Budge, Ian, Hans-Dieter Klingemann, Andrea Volkens, Judith Bara and Eric Tanenbaum. 2001. *Mapping Policy Preferences. Estimates for Parties, Electors, and Governments 1945-1998.* Oxford: Oxford University Press.

Carey, John M. 2009. *Legislative Voting and Accountability.* Cambridge: Cambridge University Press.

Clinton, Joshua, Simon Jackman and Douglas Rivers. 2004. "The Statistical Analysis of Roll Call Data." *American Political Science Review* 98 (2):355–370.

Cox, Gary W. and Mathew D. McCubbins. 2005. *Setting the Agenda: Responsible Party Government in the U.S House of Representatives.* New York: Cambridge University Press.

Desposato, Scott W. 2005. "Correcting for Small Group Inflation of Roll-Call Cohesion Scores." *B.J.Pol.S.* 35:731–744.

Duff, Andrew. 2011. "Report on a proposal for a modification of the Act concerning the election of the Members of the European Parliament by direct universal suffrage of 20 September 1976.".
URL: *http://europarl.europa.eu/sides/getDoc.do?type=REPORT&reference=A7-2011-0176&language=EN*

EP. 2012. "Your MEPs: By country and political group.".
URL: *http://www.europarl.europa.eu/meps/en/search.html*

Epstein, Leon D. 1980. *Political Parties in Western Democracies.* New Jersey: Transaction, Inc.

Faas, Thorsten. 2003. "To defect of not to defect?" *European Journal of Political Research* 42:841–866.

Farrell, David M. and Roger Scully. 2010. "The European Parliament: one parliament, several modes of political representation on the ground." *Journal of European Public Policy* 17:36–54.

Freedman, David A. 2006. "On The So-Called "Huber Sandwich Estimator" and Robust Standard Errors"." *Informal account* pp. 1–8.

Gelman, A. and J. Hill. 2007. *Data analysis using regression and multilevel/hierarchical models*. Analytical methods for social research Cambridge University Press.

Greenhill, Brian, Michael D. Ward and Audrey Sacks. 2011. "The Separation Plot: A New Visual Method for Evaluating the Fit of Binary Models." *American Journal of Political Science* 55:990–1002.

Hazan, Reuven Y. and Gideon Rahat. 2010. *Democracy within Parties*. Oxford: Oxford University Press.

Hix, S. and B. Høyland. 2011. *The Political System of the European Union*. European Union series Palgrave Macmillan.

Hix, Simon. 1997. "Executive selection in the European Union: Does the Commission president investiture procedure reduce the democratic deficit?" *European Integraion Online Papers* 1(21).

Hix, Simon. 2004. "Electoral Institutions and Legislative Behavior: Explaining Voting Defection in the European Parliament." *World Politics* 56(1):194–223.

Hix, Simon, Abdul G. Noury and Gérard Roland. 2007. *Democratic Politics in the European Parliament*. Cambridge University Press.

Hix, Simon, Abdul G. Noury and Gérard Roland. 2009. "Voting patterns and alliance formation in the European Parliament." *Phil. Trans. R. Soc.* 364:821–831.

Hix, Simon and Christopher Lord. 1996. "The Making of a President: The European Parliament and the Confirmation of Jacques Santer as President of the Commission." *Government and Opposition* 31:62–76.

Hix, Simon and Michael Marsh. 2011. "Second-order effects plus pan-European political swings: An analysis of European Parliament elections across time." *Electoral Studies* 30:4–15.

Hix, Simon, Sara Hagemann, Abdul G. Noury and Sir Julian Priestly. 2012. "Votewatch.eu.".
URL: *http://votewatch.eu*

Hooghe, Marc and Dietlind Stolle. 2005. Youth organisations within political parties: political recruitment and the transformation of party systems. In *Revisiting youth political participation*, ed. Joerg Forbrig. Council of Europe Publishing.

Jones, Mark P. 2002. Explaining the high level of party dicipline in the Argentine congress. In *Legislative Politics in Latin America*, ed. Scott Morgenstein and Benito Nacif. Cambridge University Press.

Kampstra, Peter. 2008. "Beanplot: A Boxplot Alternative of Visual Comparison of Distributions." *Journal of Statistical Software* 28:1–9.

Katz, Richard S. 2001. "The Problem of Candidate Selection and Models of Party Democracy." *Party Politics* 7:277–296.

Krehbiel, Keith. 2000. "Party Discipline and Measures of Partisanship." *American Journal of Political Science* 44:212–227.

Lam, Patrick. 2007. logit.gee: General Estimating Equation for Logistic Regression. In *Zelig: Everyone's Statistical Software*, ed. Kosuke Imai, Gary King and Olivia Lau. gking.harvard.edu/zelig.

Lehmann, Wilhelm. 2009. "The Selection of Candidates for the European Parliament by National Parties and the Impact of European Political Parties." *European Parliament* .
URL: *http://www.europarl.europa.eu/studies*

Lowe, Will, Kenneth Benoit, Slava Mikhaylov and Michael Laver. 2011. "Scaling Policy Preferences from Coded Political Texts." *Legislative Studies Quarterly* 36:123–155.

McElroy, Gail. 2006. "Committee Representation in the European Parliament." *European Union Politics* 7:5–29.

Mills, C. Wright. 2000. *The Sociological Imagination – Forthieth Anniversary Edition.* Oxford University Press.

Nordsieck, Wolfram. 2012. "Parties and Elections.".
URL: *http://parties-and-elections.de/countries.html*

Rasch, Bjørn Erik. 2000. "Floor Voting Procedures and Agenda Setting in Europe." *Legislative Studies Quarterly* 25:3–23.

Rasch, Bjorn Erik and George Tsebelis. 2011. *The Role of Governments in Legislative Agenda Setting.* Routledge/ECPR Studies in European Political Science. Routledge.

Raunio, Tapio. 2000. "Losing Independence but Gaining Recognition? Contacts Between MEPs and National Parties." *Party Politics* 6.

Rice, Stewart A. 1924. *Farmers and Workers in American Politics.* doctoral dissertation, Columbia University.

Scarrow, Susan E. 1997. "Political Career Paths and the European Parliament." *Legislative Studies Quarterly* 22(2):253–263.

Snyder, James M. Jr. and Tim Groseclose. 2000. "Estimating Party Influence on Roll-Call Voting." *American Journal of Political Science* 44:193–211.

Volden, Craig and Elizabeth Bergman. 2006. "How Strong Should Our Party By? Party Member Preferences Over Party Cohesion." *Legislative Studies Quarterly* 31:71–104.

Volkens, Andrea, Onawa Lacewell, Pola Lehmann, Sven Regel, Henrike Schultze and Annika Werner. 2011. *The Manifesto Data Collection. Manifesto Project (MRG/CMP/MARPOR).* Berlin: Wissenschaftszentrum Berlin fur Sozialforschung (WZB).

votewatch.eu. 2008. "Treaty of Lisbon (as a whole).".
URL: *http://www.votewatch.eu/cx_vote_details.php?id_act=3862&lang=en#*

votewatch.eu. 2012. "Anna Hedh (Press 'Compare to all MEPs' under 'Loyalty to political group').".
URL: *http://www.votewatch.eu/cx_mep_details.php?euro_parlamentar_id=280&lang=en&eps=0*

Yordanova, Nikoleta. 2011. "Inter-institutional Rules and Division of Power in the European Parliament: Allocation of Consultation and Co-decision Reports." *West European Politics* 34:97–121.

Zorn, Christopher J. W. 2001. "Generalized Estimating Equation Models for Correlated Data: A Review With Applications." *American Journal of Political Science* 45:470–490.

Appendix

Table 8.1: Candiate selection in European Parties

	Prop. Vote	Mem. Sys	Del. Sys
AT Grüne	1	0	1
AT FPÖ	0	0	0
AT SPÖ	0	0	0
AT ÖVP	0	0	0
BE CD&V	0	1	0
BE CDH	0	1	0
BE Ecolo	0	1	0
BE Groen	0	1	0
BE Lijst Dedecker	0	1	0
BE MR	0	1	0
BE Open VLD	0	1	0
BE SP.A	0	1	0
BE N-VA	0	0	1
BE PS	0	0	1
BE Vlaams Belang	0	0	0
BG Ataka	0	0	0
BG DPS	0	0	0
BE KB	0	0	0
BG GERB	0	0	0
CZ KDU-CSL	1	0	1
CZ CSSD	0	0	0
CZ KSCM	0	0	0
CZ ODS	0	0	0
DK SF	1	1	0
DK V	0	0	1
DK DF	1	0	1
DK FB	1	0	1
DK KF	1	0	1
DK RV	1	0	1
DK SD	1	0	1
EE PU	0	0	0
EE G	1	1	0
EE CP	1	1	0
EE RP	1	1	0
EE SDP	1	1	0
EE PPRPU	0	0	0
FI KD	0	0	0
FI KESK	0	0	0
FI VIHR	0	0	0
FI KOK	0	0	0
FI SDP	0	0	0
FI RKP	0	0	0
FR Greens	0	1	0
FR PCF	0	1	0
FR PS	0	1	0
FR FN	0	0	0
FR Libertas	0	0	0
FR MODEM	0	0	0
FR UMP	0	0	0
DE CDU	0	0	1
DE CSU	0	0	1

Source: Lehmann (2009), all errors are mine.

Table 8.1: Candidate selection: Continued.

	Prop. Vote	Mem. Sys	Del. Sys
DE FDP	0	0	1
DE Grüne	0	0	1
DE Linke	0	0	1
DE SPD	0	0	1
GR PASOK	0	1	0
GR SYRIZA	1	1	0
GR LAOS	0	0	0
GR ND	0	0	0
HU FIDESZ	0	0	0
HU SZDSZ	0	0	0
HU Jobbik	0	0	0
HU MDF	0	0	0
HU MSzP	0	0	0
IE FF	1	0	1
IE FG	1	0	1
IE Labour	1	0	1
IT IdV	0	0	0
IT LN	0	0	0
IT PD	0	0	0
IT PDL	0	0	0
IT UDC	0	0	0
LV LPP/LC	0	0	1
LV JL	0	0	0
LV PCTVL	0	0	0
LV SC	0	0	0
LV PS	0	0	0
LV TB/LNNK	0	0	0
NL GroenLinks	1	1	1
NL PvdA	0	1	1
NL CDA	0	1	0
NL SGP	0	1	0
NL D66	1	1	0
NL VVD	1	1	0
NL ChristenUnie	0	0	1
NL SP	0	0	0
PL PO	0	0	0
PL PiS	0	0	0
PL PSL	0	0	0
PL SLD-UP	0	0	0
RO SDP	0	0	0
RO PD-L	0	0	0
RO PC	0	0	0
RO PNL	0	0	0
RO UDMR	0	0	0
SK SDKÙ-DS	1	1	0
SK KDH	0	0	0
SK LS-HZDS	0	0	0
SK Smer-SD	0	0	0
SK SMK	0	0	0
SI LDS	0	0	0
SI Nsi	0	0	0
SI SD	0	0	0
SI Zares	0	0	0
ES IU-ICV	0	0	0
ES PP	0	0	0
ES PSOE	0	0	0
SW Centerparty	1	1	0
SW FP	1	1	0
SW Greens	1	1	0
SW Moderates	1	1	0
SW Left	0	0	1
SW KdS	1	0	1
SW Social Democrats	0	0	0
UK Conservatives	1	1	0
UK Greens	1	1	0
UK Labour	1	1	0
UK LibDem	1	1	0
UK BNP	0	0	0

Source: Lehmann (2009), all errors are mine.

Figure 8.1: Defection from EPG by national party

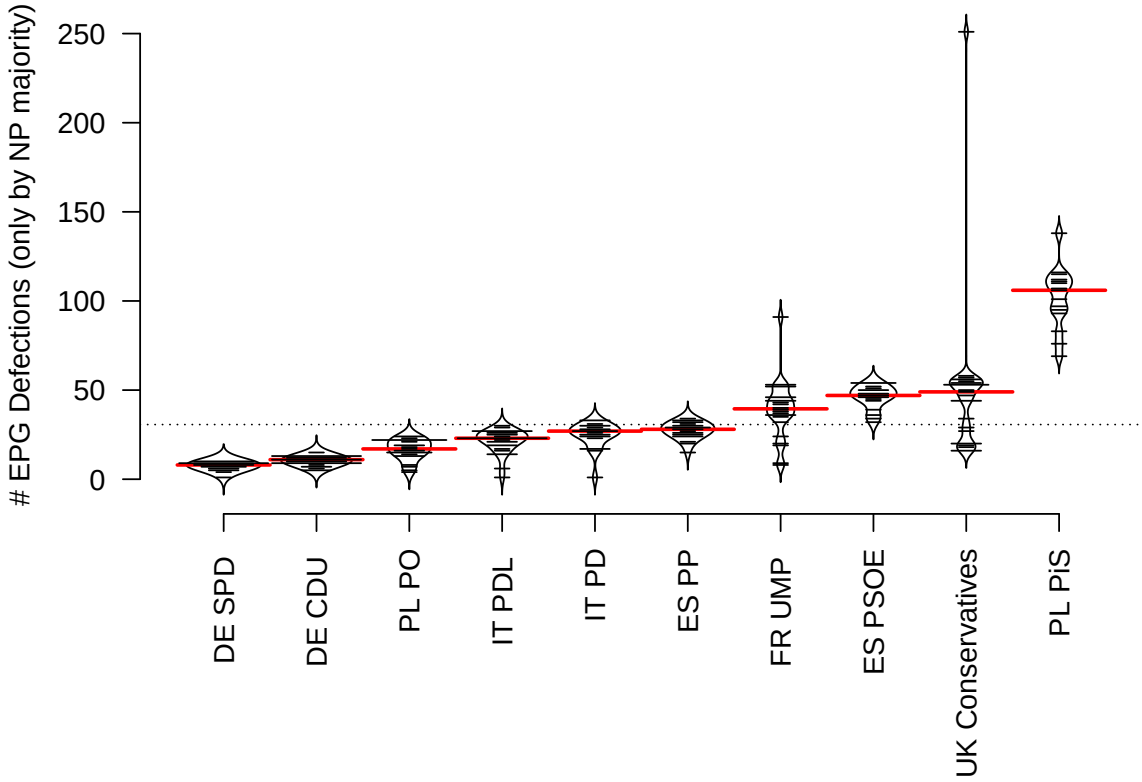


Figure 8.2: Defection from national party by MEP

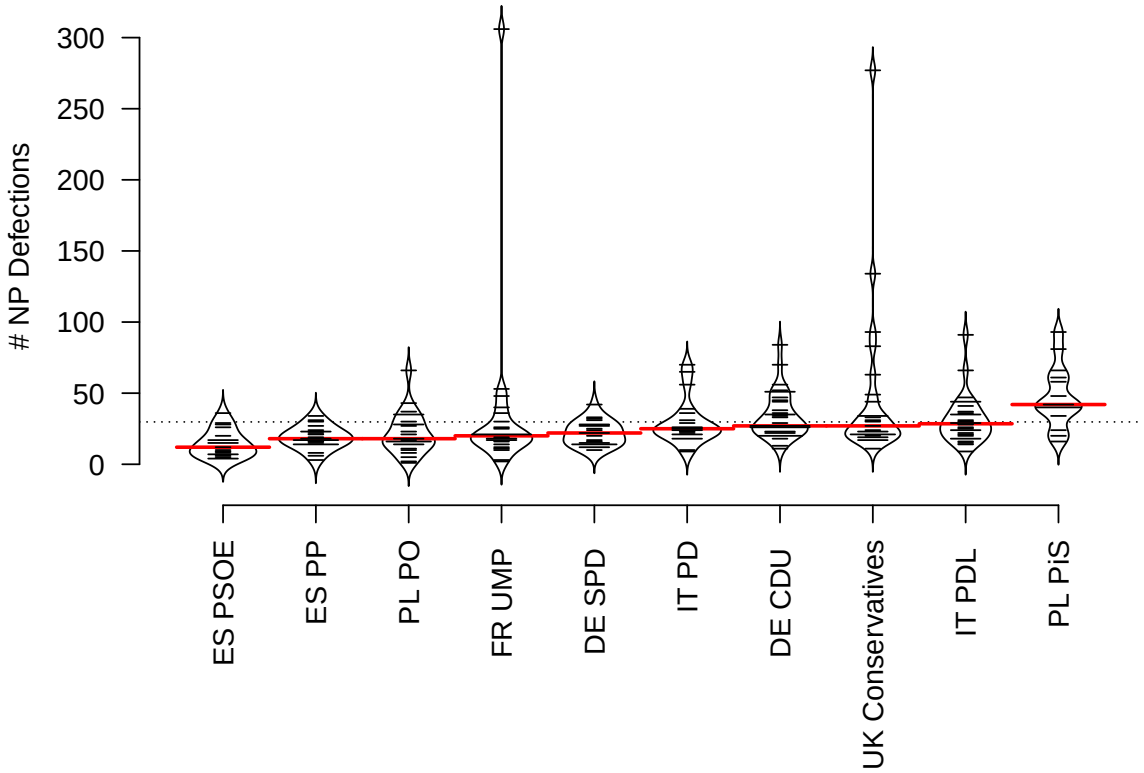


Figure 8.3: Defections, frequencies

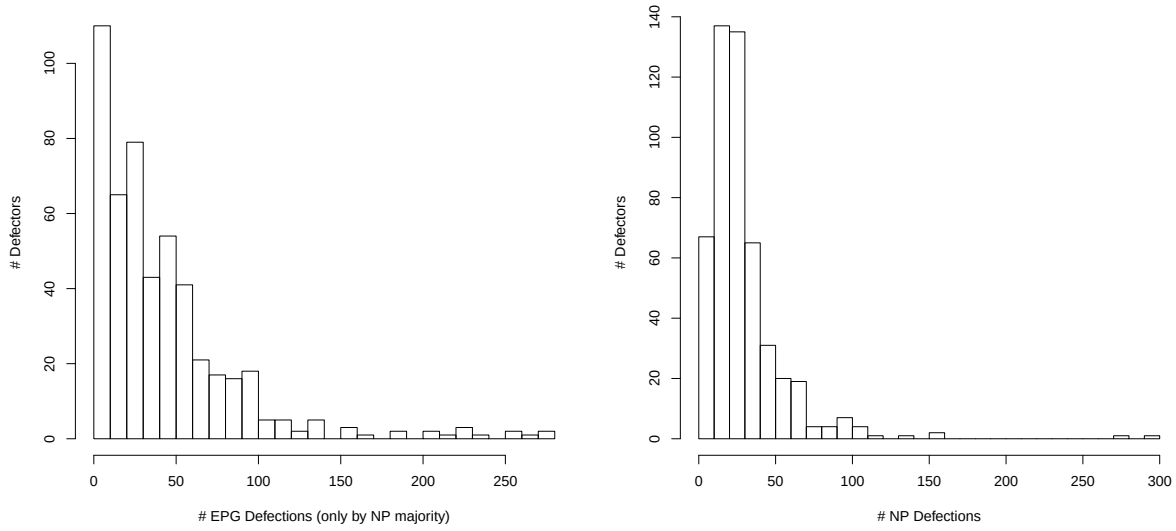


Figure 8.4: Defections in each EPG

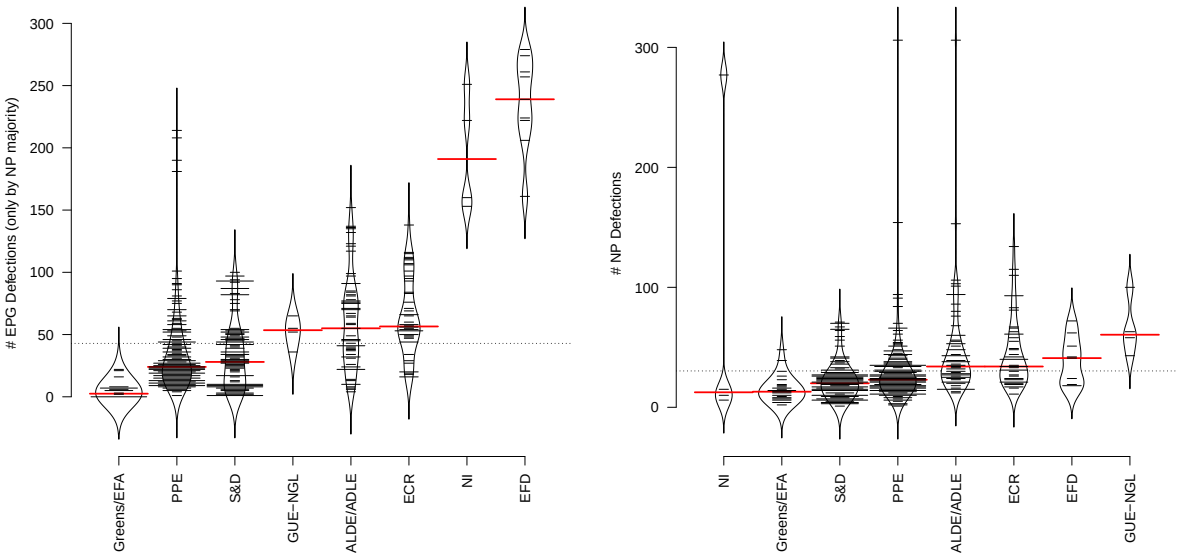


Figure 8.5: Size

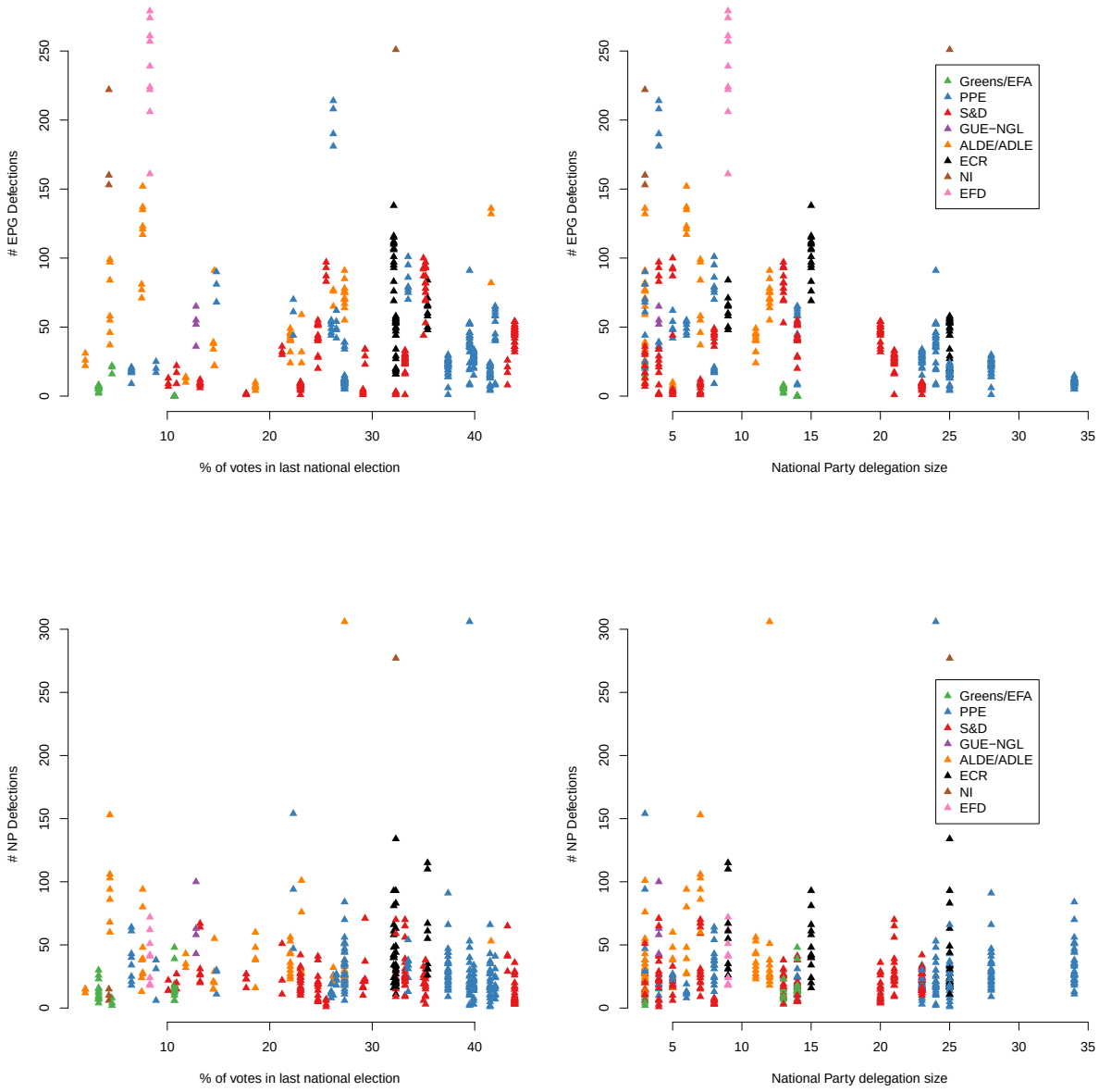


Figure 8.6: Δ Ideology

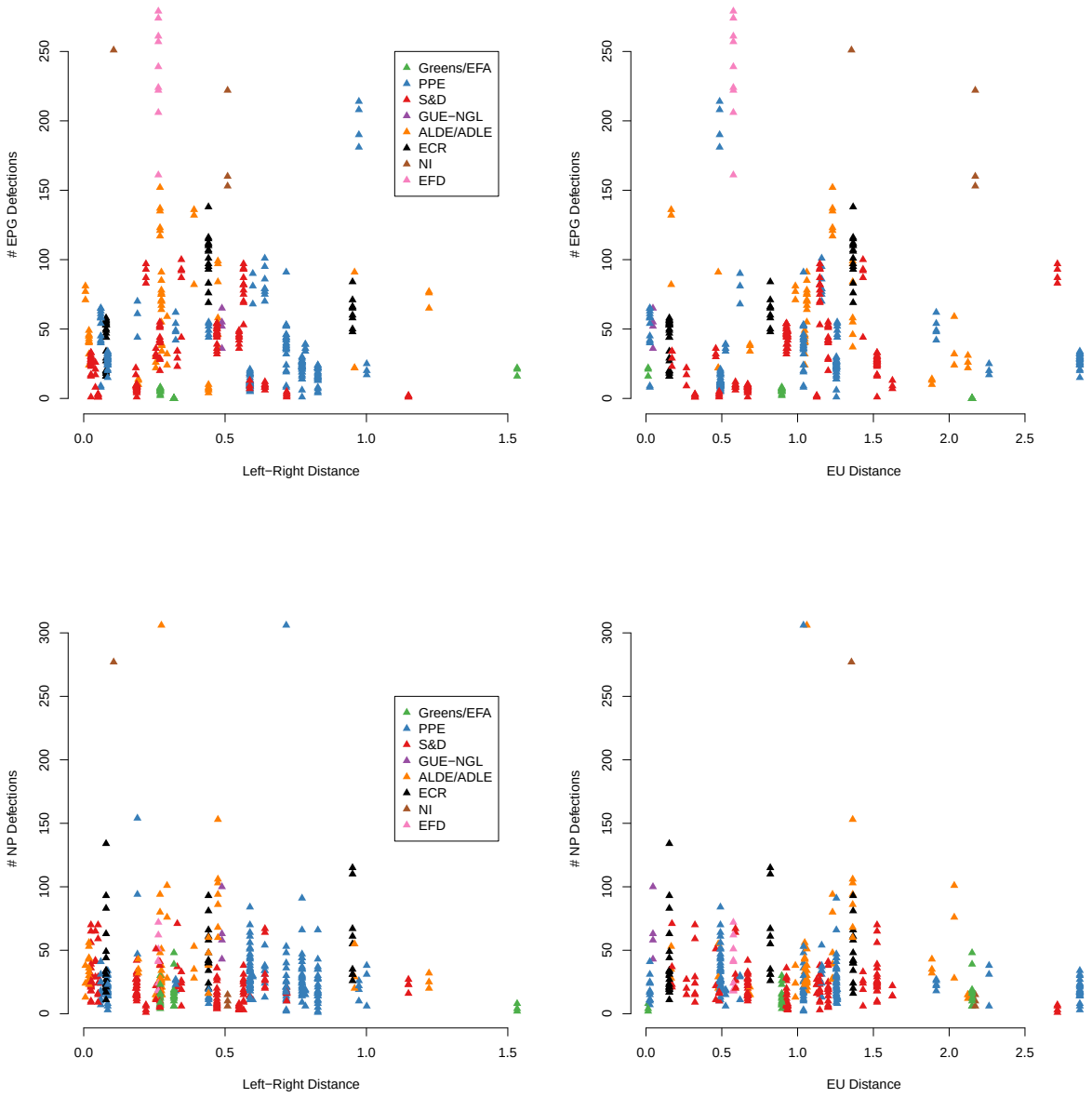


Figure 8.7: Candidate selection variables

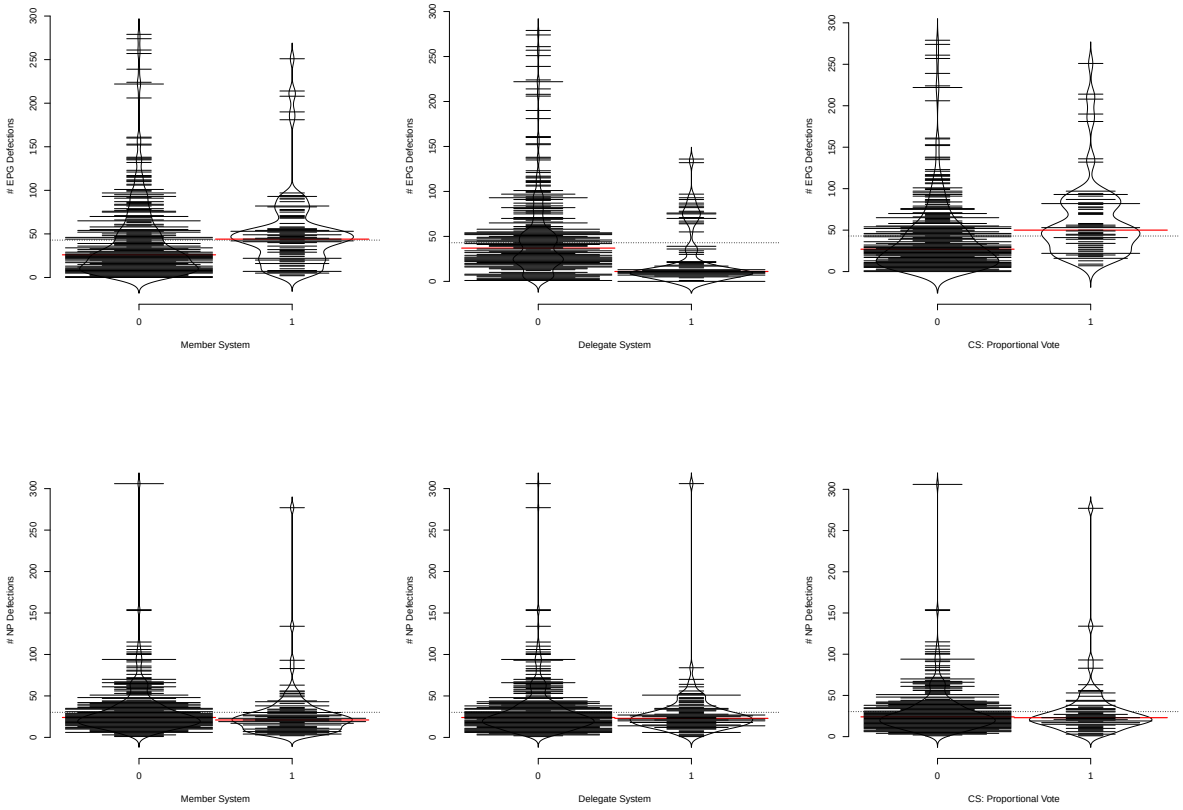


Figure 8.8: Election Variables

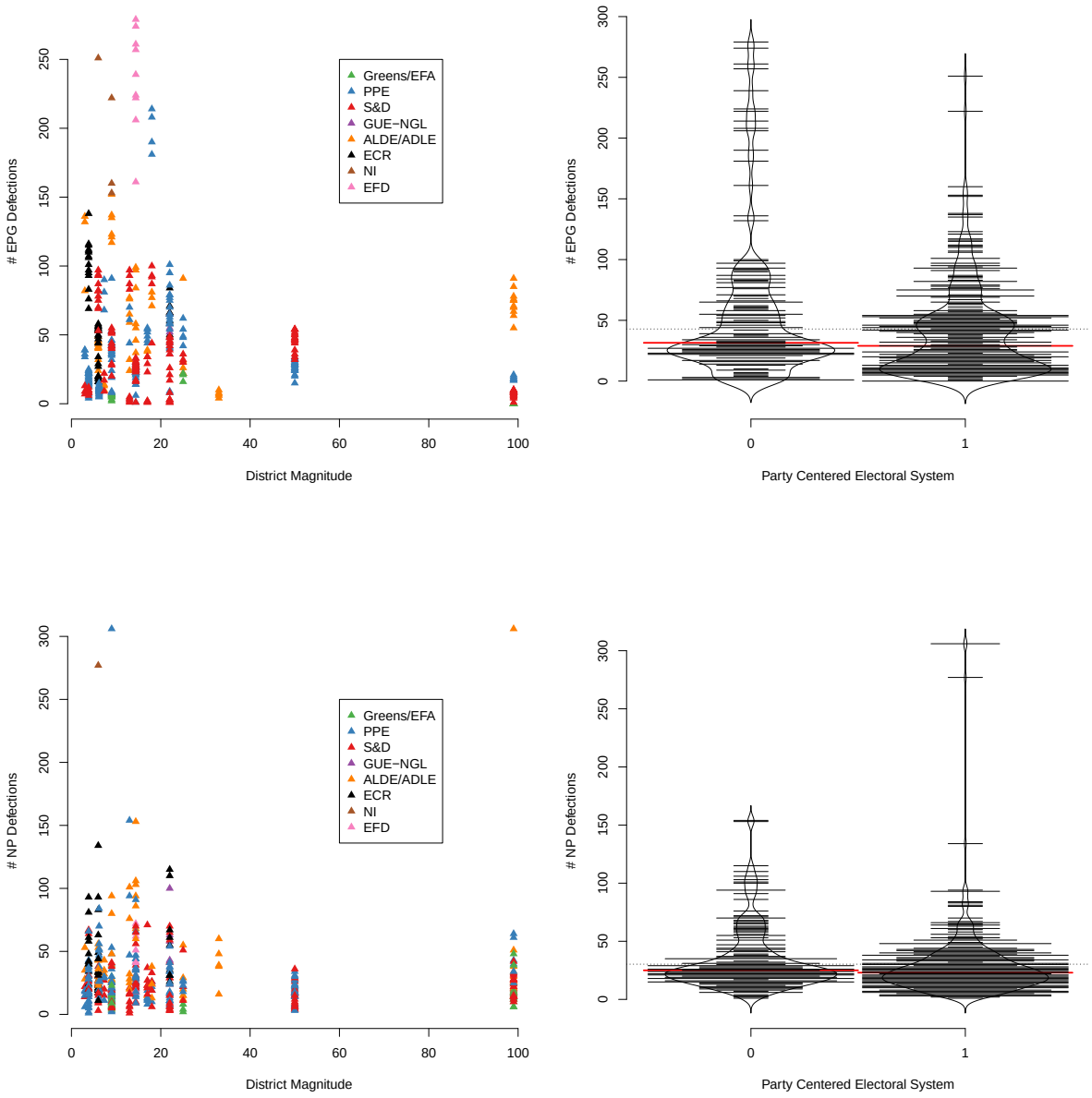
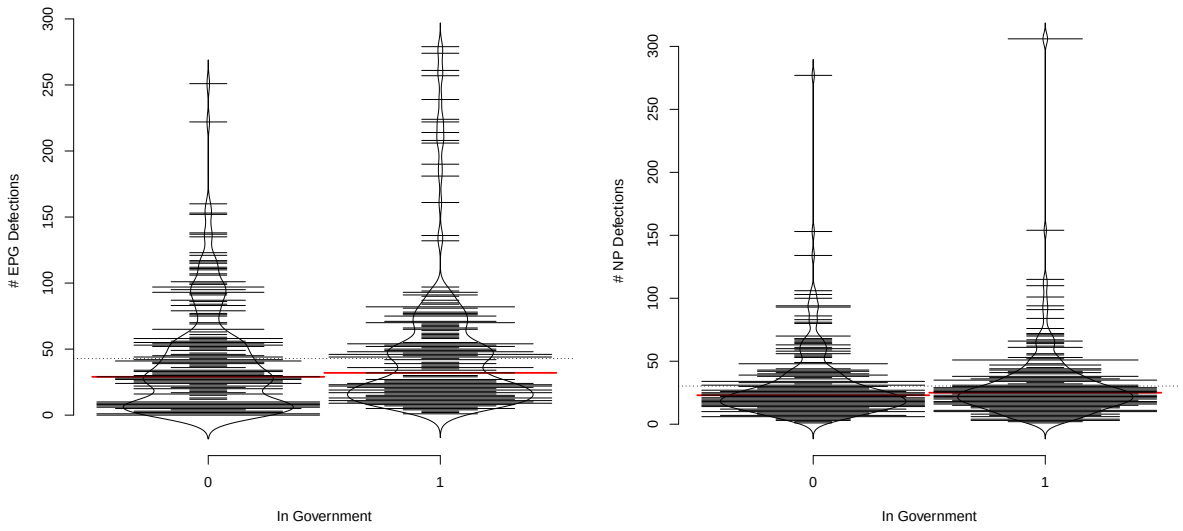


Figure 8.9: In government



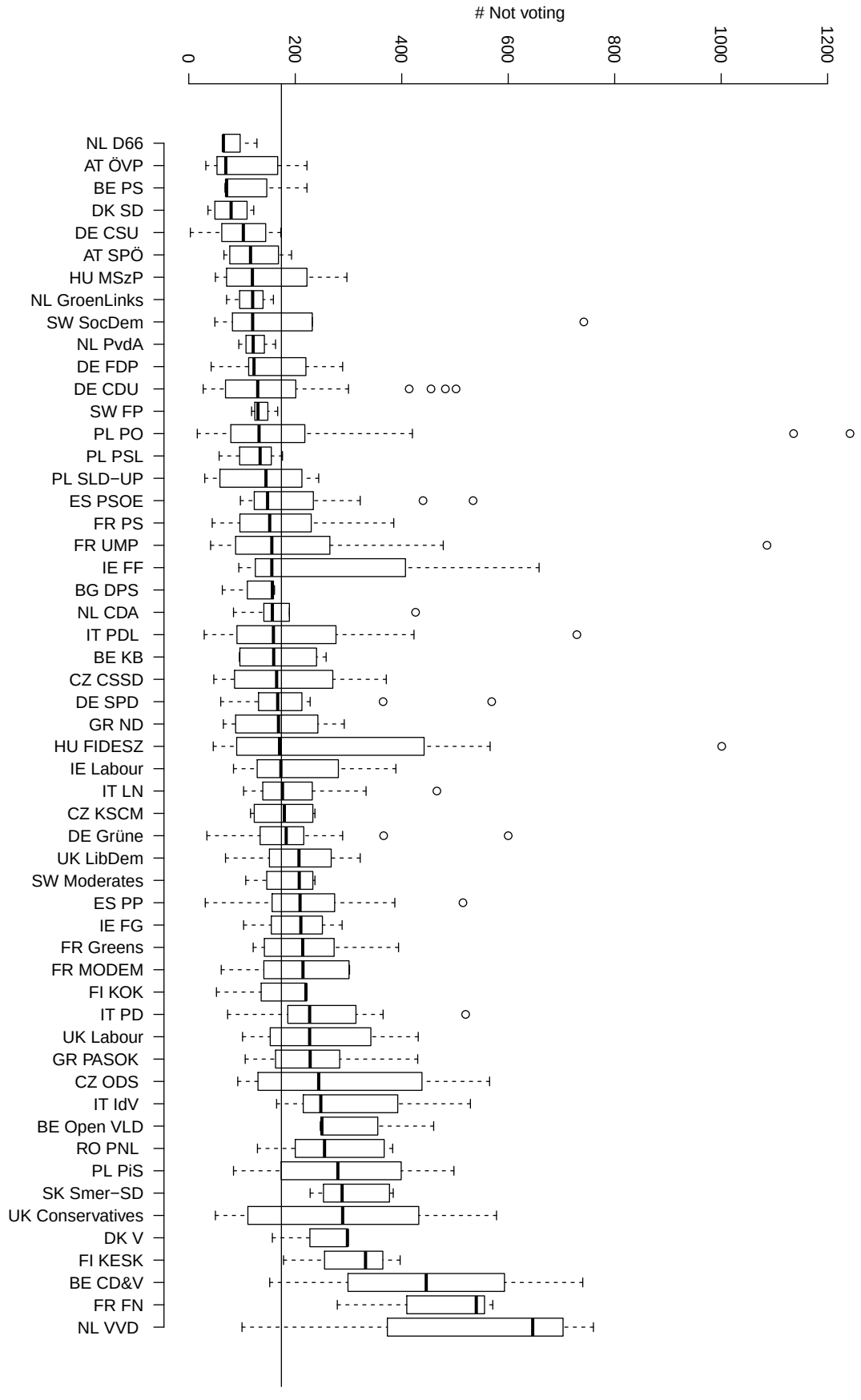


Figure 8.10: Number of non-voting occurences for each party

R-Script

```
# Produce the basic frame:
#
# Info about each MEP, their votes and a unique party-country-id.
# In addition, recodes absent to NAs.


setwd("D:/Dropbox/masteroppgave/DataConstruct/")

rm(list=ls())
library(foreign)
Votes <-  read.dta("EP7.dta")
Bio <- read.csv2("MEP_list_for_2010_survey.csv")
dim(Votes);dim(Bio)
Votes[1:10,1:10]
Bio[1:10,1:10]
Data <- merge(Bio,Votes,by.x="id",by.y="V1",all.x=TRUE)
dim(Data)
Data[1:4,1:15]

# Remove variables I do not need.
Data <- subset(Data, select=-(c(Title, EP_web_id, Full_name, E.mail.address,
                                Language.for.intro.letter)))

# Remove underscore from names. Replace with dot or V.
names(Data)[1:7] <- gsub(names(Data)[1:12],pattern="_", replacement="")
names(Data)[4:5] <- c("EPgroup", "NParty")
names(Data) <- gsub(names(Data),pattern="__", replacement="V")

# Sorting data set
Data <- Data[order(Data$NParty),]
Data <- Data[order(Data$Country),]

library(plyr)
Data$pid <- id(Data["NParty"], drop=TRUE)
Data$cid <- id(Data["Country"], drop=TRUE)
Data$pcid <- Data$cid*1000 + Data$pid

# Recoding Absent to NA. Only Yes=1, No=2 and Abstain=3 left.
for(i in 1:1351) {
  Data[,i+7] <- ifelse(Data[,i+7] == 0, NA, Data[,i+7])
}

setwd("D:/Dropbox/masteroppgave/DataConstruct/")
write.dta(Data, "Votes.dta")

remove(Bio, Votes, Data)

#########################################################################################################
setwd("D:/Dropbox/masteroppgave/rscripts/")
library(foreign)
setwd("D:/Dropbox/masteroppgave/DataConstruct/")
votes <- read.dta("Votes.dta")

# Need number of votes in subset. 1:7 and max... is all the other votes.
NumberOfVotes <- length(names(votes)[-c(1:7, (max(length(names(votes)))-2):max(length(names(votes))) )])

# Function 'mfv' in modeest calculate the most frequent value/mode.
library(modeest)
```

```
#
# # mfv
# # function (x, ...)
# # {
# #   f <- factor(x)
# #   tf <- tabulate(f)
# #   return(as.numeric(levels(f)[tf == max(tf)]))
# # }

Modes <- tapply(as.vector(votes[,8]), factor(votes$pcid), mfv)
for(i in 1:(NumberOfVotes-1)) {
  nam <- paste("Mode", i+1, sep = "")
  x <- tapply(as.vector(votes[,i+8]), factor(votes$pcid), mfv, simplify=FALSE)
  Modes <- cbind(Modes, do.call(cbind,list(x)))
  colnames(Modes)[i+1] <- nam
  colnames(Modes)[1] <- "Mode1"
}

Modes <- t(Modes)
NP <- unique(votes$pcid)

NPModes <- Modes[,1]
NPVotes <- votes[which(votes$pcid==NP[1]),]
defection <- subset(votes, votes$pcid==NP[1] ,select=c("pcid", "id"))
for(i in 1:NumberOfVotes){
  Mode <- NPModes[i]
  defection[i+2] <- ifelse( NPVotes[,i+7] %in% Mode | is.na(NPVotes[,i+7]) ,
                            ifelse(is.na(NPVotes[,i+7]), NA, 0), 1)
  nam <- paste("NPDefection", i, sep = "")
  colnames(defection)[i+2] <- nam
}
NPDef <- defection

for(z in 2:length(NP)){
  NPModes <- Modes[,z]
  NPVotes <- votes[which(votes$pcid==NP[z]),]
  defection <- subset(votes, votes$pcid==NP[z] ,select=c("pcid", "id"))
  for(i in 1:NumberOfVotes){
    Mode <- NPModes[i]
    defection[i+2] <- ifelse( NPVotes[,i+7] %in% Mode | is.na(NPVotes[,i+7]) ,
                              ifelse(is.na(NPVotes[,i+7]), NA, 0), 1)
    nam <- paste("NPDefection", i, sep = "")
    colnames(defection)[i+2] <- nam
  }
  NPDef <- rbind(NPDef, defection)
}

# Now, do the same for EP-group defections:
# Calculating the mode for all partes.
Modes <- tapply(as.vector(votes[,8]), factor(votes$EPgroup), mfv)
for(i in 1:(NumberOfVotes-1)){
  nam <- paste("Mode", i+1, sep = "")
  x <- tapply(as.vector(votes[,i+8]), votes$EPgroup, mfv, simplify=FALSE)
  Modes <- cbind(Modes, do.call(cbind,list(x)))
  colnames(Modes)[i+1] <- nam
}
colnames(Modes)[1] <- "Mode1"


Modes <- t(Modes)
EPG <- c("ALDE/ADLE",  "ECR",  "EFD",  "Greens/EFA","GUE-NGL","NI","PPE","S&D")

EPGModes <- Modes[,1]
EPGVotes <- votes[which(votes$EPgroup==EPG[1]),]
defection <- subset(votes, votes$EPgroup==EPG[1] ,select=c("pcid", "id", "EPgroup"))
for(i in 1:NumberOfVotes){
  Mode <- EPGModes[i]
  defection[i+3] <- ifelse( EPGVotes[,i+7] %in% Mode | is.na(EPGVotes[,i+7]) ,
                            ifelse(is.na(EPGVotes[,i+7]), NA, 0), 1)
  nam <- paste("EPGDefection", i, sep = "")
  colnames(defection)[i+3] <- nam
}
EPGDef <-defection
```

```
for(z in 2:length(EPG)){
  EPGModes <- Modes[,z]
  EPGVotes <- votes[which(votes$EPgroup==EPG[z]),]
  defection <- subset(votes, votes$EPgroup==EPG[z] ,select=c("pcid", "id", "EPgroup"))
  for(i in 1:NumberOfVotes){
    Mode <- EPGModes[i]
    defection[i+3] <- ifelse( EPGVotes[,i+7] %in% Mode | is.na(EPGVotes[,i+7]) ,
                              ifelse(is.na(EPGVotes[,i+7]), NA, 0), 1)
    nam <- paste("EPGDefection", i, sep = "")
    colnames(defection)[i+3] <- nam
  }
  EPGDef <- rbind(EPGDef, defection)
}



RegData <- merge(votes, EPGDef, by=c("pcid", "id", "EPgroup"))
RegData <- merge(RegData, NPDef, by=c("pcid", "id"))


setwd("D:/Dropbox/masteroppgave/DataConstruct/")
write.dta(RegData, "logitdata.dta")

#############################################################################################################

# Including the explaining variables.

setwd("D:/Dropbox/masteroppgave/DataConstruct/")

library(foreign)
RegData <- read.dta("logitdata.dta")


# Import the codes from Lehmann 2010. The set has the pcid-variable and the Manifesto-id.
lehmann <- read.csv("lehmanncodeFINAL.csv", sep=";", dec=",")
# Import election-variables.
elrule <- read.csv("elrule.csv", sep=";", dec=",")

RegData <- merge(lehmann, RegData, by=c("pcid", "Country"))
RegData <- merge(elrule, RegData, by="Country")

# Number of representatives from a country (in dataset)
countrycount <- as.data.frame(table(RegData$Country))
names(countrycount) <- c("Country", "CdelegationSize")

# Number of representatives from a party (in dataset)
pcidcount <- as.data.frame(table(RegData$pcid))
names(pcidcount) <- c("pcid", "PdelegationSize")

RegData <- merge(countrycount,RegData , by="Country")
RegData <- merge(pcidcount, RegData, by="pcid")

# CDU in Germany has different District Magnitude than the rest in Germany.
RegData$Constituencies <- ifelse(RegData$pcid==10018, 16, RegData$Constituencies)
RegData$DistrictMagnitude <- ifelse(RegData$pcid==10018, 6.1875, RegData$DistrictMagnitude)

source("d:/Dropbox/masteroppgave/rscripts/logrile.R")

RegData <- merge(CMPLogit,RegData, by.y="ManifestoID", by.x="cmp_party")

# Distance from Left-Right mean in EPgroup.
ag <- aggregate(logrile ~ EPgroup, data=RegData, mean)
names(ag) <- c("EPgroup", "lrmean")
RegData <- merge(ag, RegData, by="EPgroup")
RegData$PrefDiff <- abs(RegData$logrile - RegData$lrmean)

# Distance from EU integration mean in EPgroup
ag <- aggregate(logeu ~ EPgroup, data=RegData, mean)
names(ag) <- c("EPgroup", "eumean")
RegData <- merge(ag, RegData, by="EPgroup")
```

```
RegData$EUPrefDiff <- abs(RegData$logeu - RegData$eumean)


# Remove parties with 2 or less delegates in data set
RegData <- subset(RegData, RegData$PdelegationSize>2)

remove(CMPLogit, ag, countrycount, elrule, lehmann, pcidcount)

# Reshape the dataset by taking all votes into one variable, copying the information about the MEP who voted.
library(reshape)
RegData <- melt(RegData, id.vars=c("EPgroup","id","Country","pcid","Representatives","Constituencies",
                                   "DistrictMagnitude","PrefVote","PropRule","Threshold","ThreshBin","Party",
                                   "NationalParty","CountryParty","PropVote","MemSys","DelSys","VetoElite","NElecRes",
                                   "Gov","CdelegationSize","PdelegationSize","cmp_edate","logeu","logrile","Lastname",
                                   "Firstname","NParty","lrmean","PrefDiff","eumean","EUPrefDiff",
                                   "cmp_party", "VoteID", "pid", "cid"
                                   ))

# Three types of information from the votes: Defection from NP, from EP and the vote itself.
# Just cutting and pasting them together again.
RegData2 <- RegData[which(RegData$variable %in% unique(RegData$variable)[2703:4053]),]
RegData3 <- RegData[which(RegData$variable %in% unique(RegData$variable)[1352:2702]),]
RegData4 <- RegData[which(RegData$variable %in% unique(RegData$variable)[1:1351]),]

RegData <- cbind(RegData2, RegData3[37:38])
RegData <- cbind(RegData, RegData4[37:38])

remove(RegData2, RegData3, RegData4)

names(RegData)[37] <- "ID.NPDefection"
names(RegData)[38] <- "NPDefection"
names(RegData)[39] <- "ID.EPGDefection"
names(RegData)[40] <- "EPGDefection"
names(RegData)[41] <- "ID.Vote"
names(RegData)[42] <- "Voted"

# Voted.csv is information about each vote.
setwd("D:/Dropbox/masteroppgave/DataConstruct")
voteinfo <- read.table("voted.csv", sep=";", header=TRUE)
# I have only the first 1351 votes.
voteinfo <- voteinfo[1:1351,]

# Make VoteID similar to name of votes.
voteinfo$VoteID <- paste("NPDefection", voteinfo$VoteID, sep="")

RegData <- merge(RegData, voteinfo, by.x="ID.NPDefection", by.y="VoteID")

RegData$VoteDiff <- abs(RegData$Yeas - RegData$No)
RegData$CloseVote <- ifelse(RegData$VoteDiff<=
  0.1*(RegData$Yeas + RegData$No), 1, 0)

# Produce variable that Hix proposes.
RegData$partycentered <- ifelse(RegData$PrefVote==0 & RegData$PropRule==1, 1,0)

# Make factor variables factors.
RegData$PropVote <- with(RegData, factor(PropVote))
RegData$MemSys <- with(RegData, factor(MemSys))
RegData$DelSys <- with(RegData, factor(DelSys))
RegData$VetoElite <- with(RegData, factor(VetoElite))
RegData$PrefVote <- with(RegData, factor(PrefVote))
RegData$PropRule <- with(RegData, factor(PropRule))
RegData$ThreshBin <- with(RegData, factor(ThreshBin))
RegData$partycentered <- with(RegData, factor(partycentered))
RegData$EPgroup <- with(RegData, factor(EPgroup))
RegData$Country <- with(RegData, factor(Country))
RegData$Gov <- with(RegData, factor(Gov))

RegData$DistrictMagnitude <- as.numeric(RegData$DistrictMagnitude)

remove(voteinfo)

#############################################################################################################
```

```
##### Graphical analysis #####

setwd("D:/Dropbox/masteroppgave/DataConstruct/")
write.dta(RegData, "RegData.dta")

setwd("D:/Dropbox/masteroppgave/rscripts/")
source("logitPrep.R")

# Counting number of defections for each MEP
RegData$OnlyEPDef <- ifelse(RegData$EPGDefection==1 & RegData$NPDefection==0, 1, 0)
OnlyEPDefNum <- tapply(RegData$OnlyEPDef, RegData$id, sum, na.rm=TRUE)
OnlyEPDefNum <- as.data.frame(OnlyEPDefNum)
OnlyEPDefNum$id <- rownames(OnlyEPDefNum)
RegData <- merge(OnlyEPDefNum, RegData, by="id")

NPDefNum <- tapply(RegData$NPDefection, RegData$id, sum, na.rm=TRUE)
NPDefNum <- as.data.frame(NPDefNum)
NPDefNum$id <- rownames(NPDefNum)
RegData <- merge(NPDefNum, RegData, by="id")

# Abstention
library(gmodels)
CrossTable(RegData$NPDefection, (RegData$Voted==3), digits=2, expected=FALSE,
           prop.r=TRUE, prop.c=TRUE, prop.t=TRUE, prop.chisq=FALSE, format="SPSS")


CrossTable((RegData$EPGDefection==1 & RegData$NPDefection==0), (RegData$Voted==3),
           digits=2, expected=FALSE,
           prop.r=TRUE, prop.c=TRUE, prop.t=TRUE, prop.chisq=FALSE, format="SPSS")

# Remove all data which changes for each vote.
RegData2 <- unique(RegData[-c(4,40:66,68:69)])


library(beanplot)
# Use a different palette, chosen by RColorBrewer, with minor changes.
library(RColorBrewer)
mypalette <- brewer.pal(8, "Set1")
mypalette <- c("#4DAF4A", "#377EB8", "#E41A1C", "#984EA3", "#FF7F00", "black", "#A65628", "#F781BF")

setwd("d:/Dropbox/masteroppgave/graphs/")

pdf(file="EPGOnlyEPDefNumhist.pdf")
hist(RegData2$OnlyEPDefNum, breaks=30, main="",
     xlab="# EPG Defections (only by NP majority)", ylab="# Defectors")
dev.off()

# Fix the ordering
oi <- order(tapply(RegData2$OnlyEPDefNum, RegData2$EPgroup, median))
RegData2$EPgroup <- factor(RegData2$EPgroup, levels(RegData2$EPgroup)[oi])

pdf(file="OnlyEPDefNum.pdf")
par(mar=c(6, 4, 2, 2), bty="n")
RegData2$EPgroup <- with(RegData2, factor(EPgroup))
beanplot(OnlyEPDefNum~EPgroup, data = RegData2,
         ylab="# EPG Defections (only by NP majority)", las=2, log="",
         col=c("white", "black", "black", "red"), beanlines="median",
         overalline="median")
dev.off()

# Beanplot only work with a certain amount of data.
x <- as.data.frame(ifelse(table(RegData2$CountryParty)>=15,1,0))
x <- subset(x, x[1]==1)
RegData3 <- RegData2[which(RegData2$CountryParty %in% row.names(x)),]

# Fix the ordering
oi <- order(tapply(RegData3$OnlyEPDefNum, RegData3$CountryParty, median))
RegData3$CountryParty <- factor(RegData3$CountryParty, levels(RegData3$CountryParty)[oi])


pdf(file="OnlyEPDefNumNP.pdf", width=7, height=5)
par(mar=c(8, 4, 2, 2), bty="n")
RegData3$CountryParty <- with(RegData3, factor(CountryParty))
```

```
beanplot(OnlyEPDefNum ~ CountryParty, data = RegData3,
         ylab="# EPG Defections (only by NP majority)", las=2, log="",
         col=c("white", "black", "black", "red"), beanlines="median",overalline="median")
dev.off()


pdf(file="EPGlogrile.pdf")
par(mar=c(5, 4, 4, 2), bty="n")
with(RegData2, plot(logrile, OnlyEPDefNum, xlab="Left-Right", ylab="# EPG Defections" , pch=17,
                    col=mypalette[RegData2$EPgroup]))
legend(-1,250,col = mypalette, legend=levels(RegData2$EPgroup), pch=17, cex=1)
dev.off()

pdf(file="EPGPrefDiff.pdf")
par(mar=c(5, 4, 4, 2), bty="n")
with(RegData2, plot(PrefDiff, OnlyEPDefNum, xlab="Left-Right Distance", ylab="# EPG Defections" , pch=17,
                    col=mypalette[RegData2$EPgroup]))
legend(1.1,250,col = mypalette, legend=levels(RegData2$EPgroup), pch=17, cex=1)
dev.off()

pdf(file="EPGlogeu.pdf")
par(mar=c(5, 4, 4, 2), bty="n")
with(RegData2, plot(logeu, OnlyEPDefNum, xlab="EU Integration", ylab="# EPG Defections" , pch=17,
                    col=mypalette[RegData2$EPgroup]))
dev.off()

pdf(file="EPGEUPrefDiff.pdf")
par(mar=c(5, 4, 4, 2), bty="n")
with(RegData2, plot(EUPrefDiff, OnlyEPDefNum, xlab="EU Distance", ylab="# EPG Defections" , pch=17,
                    col=mypalette[RegData2$EPgroup]))
dev.off()


pdf(file="EPGnelecres.pdf")
par(mar=c(5, 4, 4, 2), bty="n")
with(RegData2, plot(NElecRes, OnlyEPDefNum, xlab="% of votes in last national election", ylab="# EPG Defections" , pch=17,
                    col=mypalette[RegData2$EPgroup]))
dev.off()

pdf(file="EPGPdelegationSize.pdf")
par(mar=c(5, 4, 4, 2), bty="n")
with(RegData2, plot(PdelegationSize, OnlyEPDefNum, xlab="National Party delegation size", ylab="# EPG Defections" , pch=17,
                    col=mypalette[RegData2$EPgroup]))
legend(24,240,col = mypalette, legend=levels(RegData2$EPgroup), pch=17, cex=1)
dev.off()

pdf("EPGMemSys.pdf")
par(mar=c(5, 4, 4, 2), bty="n")
beanplot(OnlyEPDefNum ~ MemSys, data = RegData2,xlab="Member System", ylab="# EPG Defections",
         col=c("white", "black", "black", "red"), beanlines="median", log="",
         overalline="median")
dev.off()

pdf("EPGDelSys.pdf")
par(mar=c(5, 4, 4, 2), bty="n")
beanplot(OnlyEPDefNum ~ DelSys, data = RegData2,xlab="Delegate System", ylab="# EPG Defections",
        col=c("white", "black", "black", "red"), beanlines="median", log="",
        overalline="median")
dev.off()


pdf("EPGPropVote.pdf")
par(mar=c(5, 4, 4, 2), bty="n")
beanplot(OnlyEPDefNum ~ PropVote, data = RegData2,xlab="CS: Proportional Vote",
        ylab="# EPG Defections",col=c("white", "black", "black", "red"), beanlines="median",
        overalline="median", log="")
dev.off()

pdf("EPGDistrictMagnitude.pdf")
par(mar=c(5, 4, 4, 2), bty="n")
with(RegData2, plot(as.numeric(DistrictMagnitude), OnlyEPDefNum,
                   xlab="District Magnitude", ylab="# EPG Defections", pch=17,
                    col=mypalette[RegData2$EPgroup]))
```

```
legend(60,250,col = mypalette, legend=levels(RegData2$EPgroup), pch=17, cex=1)
dev.off()

pdf("EPGpartycentered.pdf")
par(mar=c(5, 4, 4, 2), bty="n")
beanplot(OnlyEPDefNum ~ partycentered, data = RegData2, xlab="Party Centered Electoral System",
        ylab="# EPG Defections",col=c("white", "black", "black", "red"), beanlines="median",
        overalline="median", log="")
dev.off()

pdf("EPGGov.pdf")
par(mar=c(5, 4, 4, 2), bty="n")
beanplot(OnlyEPDefNum ~ Gov, data = RegData2, xlab="In Government", ylab="# EPG Defections",
        col=c("white", "black", "black", "red"), beanlines="median",
        overalline="median", log="")
dev.off()

##### Then the same for national party defection #####

pdf(file="NPDefNumhist.pdf")
hist(RegData2$NPDefNum, breaks=30, main="",
     xlab="# NP Defections", ylab="# Defectors")
dev.off()

# Fix the ordering
oi <- order(tapply(RegData2$NPDefNum, RegData2$EPgroup, median))
RegData2$EPgroup <- factor(RegData2$EPgroup, levels(RegData2$EPgroup)[oi])

pdf(file="NPDefNum.pdf")
par(mar=c(6, 4, 2, 2), bty="n")
RegData2$EPgroup <- with(RegData2, factor(EPgroup))
beanplot(NPDefNum~EPgroup, data = RegData2,
         ylab="# NP Defections", las=2, log="",
         col=c("white", "black", "black", "red"), beanlines="median",
         overalline="median")
dev.off()



# Beanplot only work with a certain amount of data.
x <- as.data.frame(ifelse(table(RegData2$CountryParty)>=15,1,0))
x <- subset(x, x[1]==1)
RegData3 <- RegData2[which(RegData2$CountryParty %in% row.names(x)),]

# Fix the ordering
oi <- order(tapply(RegData3$NPDefNum, RegData3$CountryParty, median))
RegData3$CountryParty <- factor(RegData3$CountryParty, levels(RegData3$CountryParty)[oi])


pdf(file="NPDefNumNP.pdf", width=7, height=5)
par(mar=c(8, 4, 2, 2), bty="n")
RegData3$CountryParty <- with(RegData3, factor(CountryParty))
beanplot(NPDefNum ~ CountryParty, data = RegData3,
         ylab="# NP Defections", las=2, log="",
         col=c("white", "black", "black", "red"), beanlines="median",overalline="median")
dev.off()

# Fix the ordering (for colors)
oi <- order(tapply(RegData2$OnlyEPDefNum, RegData2$EPgroup, median))
RegData2$EPgroup <- factor(RegData2$EPgroup, levels(RegData2$EPgroup)[oi])


pdf(file="NPlogrile.pdf")
par(mar=c(5, 4, 4, 2), bty="n")
with(RegData2, plot(logrile, NPDefNum, xlab="Left-Right", ylab="# NP Defections" , pch=17,
                    col=mypalette[RegData2$EPgroup]))
legend(-1,250,col = mypalette, legend=levels(RegData2$EPgroup), pch=17, cex=1)
dev.off()

pdf(file="NPPrefDiff.pdf")
par(mar=c(5, 4, 4, 2), bty="n")
with(RegData2, plot(PrefDiff, NPDefNum, xlab="Left-Right Distance", ylab="# NP Defections" , pch=17,
                    col=mypalette[RegData2$EPgroup]))
```

```
legend(1.1,250,col = mypalette, legend=levels(RegData2$EPgroup), pch=17, cex=1)
dev.off()

pdf(file="NPlogeu.pdf")
par(mar=c(5, 4, 4, 2), bty="n")
with(RegData2, plot(logeu, NPDefNum, xlab="EU Integration", ylab="# NP Defections" , pch=17,
                    col=mypalette[RegData2$EPgroup]))
dev.off()

pdf(file="NPEUPrefDiff.pdf")
par(mar=c(5, 4, 4, 2), bty="n")
with(RegData2, plot(EUPrefDiff, NPDefNum, xlab="EU Distance", ylab="# NP Defections" , pch=17,
                    col=mypalette[RegData2$EPgroup]))
dev.off()


pdf(file="NPnelecres.pdf")
par(mar=c(5, 4, 4, 2), bty="n")
with(RegData2, plot(NElecRes, NPDefNum, xlab="% of votes in last national election", ylab="# NP Defections" , pch=17,
                    col=mypalette[RegData2$EPgroup]))
dev.off()

pdf(file="NPPdelegationSize.pdf")
par(mar=c(5, 4, 4, 2), bty="n")
with(RegData2, plot(PdelegationSize, NPDefNum, xlab="National Party delegation size", ylab="# NP Defections" , pch=17,
                    col=mypalette[RegData2$EPgroup]))
legend(24,260,col = mypalette, legend=levels(RegData2$EPgroup), pch=17, cex=1)
dev.off()

pdf("NPMemSys.pdf")
par(mar=c(5, 4, 4, 2), bty="n")
beanplot(NPDefNum ~ MemSys, data = RegData2,xlab="Member System", ylab="# NP Defections",
         col=c("white", "black", "black", "red"), beanlines="median",
         overalline="median", log="")
dev.off()

pdf("NPDelSys.pdf")
par(mar=c(5, 4, 4, 2), bty="n")
beanplot(NPDefNum ~ DelSys, data = RegData2,xlab="Delegate System", ylab="# NP Defections",
         col=c("white", "black", "black", "red"), beanlines="median",
         overalline="median", log="")
dev.off()


pdf("NPPropVote.pdf")
par(mar=c(5, 4, 4, 2), bty="n")
beanplot(NPDefNum ~ PropVote, data = RegData2,xlab="CS: Proportional Vote",
         ylab="# NP Defections",col=c("white", "black", "black", "red"), beanlines="median",
         overalline="median", log="")
dev.off()

pdf("NPDistrictMagnitude.pdf")
par(mar=c(5, 4, 4, 2), bty="n")
with(RegData2, plot(as.numeric(DistrictMagnitude), NPDefNum,
                    xlab="District Magnitude", ylab="# NP Defections", pch=17,
                    col=mypalette[RegData2$EPgroup]))
legend(60,250,col = mypalette, legend=levels(RegData2$EPgroup), pch=17, cex=1)
dev.off()

pdf("NPpartycentered.pdf")
par(mar=c(5, 4, 4, 2), bty="n")
beanplot(NPDefNum ~ partycentered, data = RegData2, xlab="Party Centered Electoral System",
         ylab="# NP Defections",col=c("white", "black", "black", "red"), beanlines="median",
         overalline="median", log="")
dev.off()

pdf("NPGov.pdf")
par(mar=c(5, 4, 4, 2), bty="n")
beanplot(NPDefNum ~ Gov, data = RegData2, xlab="In Government", ylab="# NP Defections",
         col=c("white", "black", "black", "red"), beanlines="median",
         overalline="median", log="")
dev.off()
```

```
#################################################################################################

##### REGRESSION MODELS #####

setwd("D:/Dropbox/masteroppgave/rscripts/")
source("logitPrep.R")

library(car)
library(Zelig)
library(xtable)
library(arm)
library(lme4)

# CrossTable for number of defections
library(gmodels)
CrossTable(RegData$NPDefection, RegData$EPGDefection, digits=2, expected=FALSE,
           prop.r=TRUE, prop.c=TRUE, prop.t=TRUE, prop.chisq=FALSE, format="SPSS")

# Only one party in GUE-NGL and EFD in dataset, two in NI.
# For consistence over models, I choose to leave these out of my models.
RegData <- RegData[which(!(RegData$EPgroup=="GUE-NGL"|RegData$EPgroup=="EFD"|RegData$EPgroup=="NI")),]

RegData$EPgroup <- factor(RegData$EPgroup)

# Centering variables for more senseful interpretation.
RegData$c.PdelegationSize <- RegData$PdelegationSize - mean(RegData$PdelegationSize)
RegData$c.NElecRes <- RegData$NElecRes - mean(RegData$NElecRes)
RegData$c.DistrictMagnitude <- RegData$DistrictMagnitude - mean(RegData$DistrictMagnitude)


# First, naive model.

logitModel <- glm((EPGDefection == 1 & NPDefection == 0) ~ CloseVote + FinalVote + (Type=="Leg") +
  PrefDiff + EUPrefDiff + c.PdelegationSize + c.NElecRes + Gov + MemSys + DelSys + partycentered +
  c.DistrictMagnitude + EPgroup,
                  data=RegData[which(RegData$VotingRule=="s"),], family=binomial(link = "logit"))
summary(logitModel)

setwd("D:/Dropbox/masteroppgave/tex/")
sink("logitModel.txt")
summary(logitModel)
str(logitModel)
logLik(logitModel)
sink()

# Multi-level models

M1 <- lmer((EPGDefection == 1 & NPDefection == 0) ~ 1 + (1|EPgroup) +  (1|ID.EPGDefection),
           data=RegData[which(RegData$VotingRule=="s"),], family=binomial(link = "logit"))

setwd("D:/Dropbox/masteroppgave/tex/")
sink("M1.txt")
summary(M1)
str(M1)
sink()

M1Coef <- coef(M1)$ID.EPGDefection
M1SE <- se.coef(M1)$ID.EPGDefection

M1Results <- cbind(M1Coef, M1SE)
M1Results <- M1Results[order(M1Results[1]),]

setwd("D:/Dropbox/masteroppgave/graphs/")
pdf("ExpectedProbM1.pdf", width=7, height=5)
plot (1, xlim=c(0,1223), ylim=range(max(invlogit(M1Results[1]+M1Results[2]))+0.05,
                                   min(invlogit(M1Results[1]-M1Results[2]))),
      pch=20, ylab="E(Pr) for average group", xlab="Vote", bty="n")

for (i in 1:1223){
  points(i, invlogit(M1Results[i,1]), pch=".")
  lines (rep(i,2), invlogit(M1Results[i,1]+M1Results[i,2]*c(-2,2)), lwd=.5, col="lightblue")
}
dev.off()
```

```
# Make squared version of Delegation Size-variable.
RegData$PdelegationSize.sq <- RegData$PdelegationSize*RegData$PdelegationSize

M2 <- lmer((EPGDefection == 1 & NPDefection == 0) ~ 1 + CloseVote + FinalVote + (Type=="Leg") +
  PrefDiff + EUPrefDiff + PdelegationSize + PdelegationSize.sq + c.NElecRes + Gov +(1|EPgroup) +  (1|ID.EPGDefection),
          data=RegData[which(RegData$VotingRule=="s"),], family=binomial(link = "logit"))

sink("M2sq.txt")
summary(M2)
str(M2)
sink()


beta <- fixef(M2)
covvar <- vcov(M2)

x <- cbind(1,0,0,0,0.5,0,seq(3,34,1),(seq(3,34,1)*seq(3,34,1)), 0, 0)

beta.sim <- mvrnorm(n=1000, beta,covvar)
xb <- x %*% t(beta.sim)
p <- 1/(1 + exp(-xb))

predicted.prob <- t(apply(p,1,quantile, probs = c(.025,.5,.975)))
cord.y <- c(predicted.prob[,1],rev(predicted.prob[,3]))
cord.x <- c(3:34,34:3)

setwd("D:/Dropbox/masteroppgave/graphs/")
pdf("predDelegationSize.pdf", width=10, height=7)
plot(0,0, type="n",xlim=c(3,34),
     ylim=c(0,.05), main="Delegation Size on defection by NP",
     xlab="Number of delegates from national party",
     ylab="Probability of defection by national party",bty="n",
     sub="")
polygon(cord.x,cord.y,col="lightblue",border=NA)
lines(x[,7],predicted.prob[,2],lty=1) # estimated effect
dev.off()


M8 <- lmer(NPDefection ~ 1 + CloseVote + FinalVote + (Type=="Leg") +
  c.PdelegationSize + c.NElecRes + Gov + MemSys + DelSys + partycentered + c.DistrictMagnitude +
  (1|EPgroup) +  (1|ID.EPGDefection),
          data=RegData[which(RegData$VotingRule=="s"),], family=binomial(link = "logit"))

setwd("D:/Dropbox/masteroppgave/tex/")
sink("M8.txt")
summary(M8)
str(M8)
sink()

M8PropVote <- lmer(NPDefection ~ 1 + CloseVote + FinalVote + (Type=="Leg") +
  c.PdelegationSize + c.NElecRes + Gov +PropVote + partycentered + c.DistrictMagnitude +
  (1|EPgroup) +  (1|ID.EPGDefection),
          data=RegData[which(RegData$VotingRule=="s"),], family=binomial(link = "logit"))

sink("M8PropVote.txt")
summary(M8PropVote)
str(M8PropVote)
sink()


M9 <- lmer(NPDefection ~ 1 + (1|EPgroup) +  (1|ID.EPGDefection),
        data=RegData[which(RegData$VotingRule=="s"),], family=binomial(link = "logit"))

sink("M9.txt")
summary(M9)
str(M9)
sink()

# Varying slope & intercept models

M4 <- lmer((EPGDefection == 1 & NPDefection == 0) ~ 1 + FinalVote +
  (1+FinalVote|EPgroup) +  (1|ID.EPGDefection),
```

```
          data=RegData, family=binomial(link = "logit"))

sink("M4.txt")
summary(M4)
str(M4)
coef(M4)$EPgroup
se.coef(M4)$EPgroup
sink()

M5 <- lmer((EPGDefection == 1 & NPDefection == 0) ~ 1 + (Type=="Leg") +
  (1+(Type=="Leg")|EPgroup) +  (1|ID.EPGDefection),
          data=RegData, family=binomial(link = "logit"))

sink("M5.txt")
summary(M5)
str(M5)
coef(M5)$EPgroup
se.coef(M5)$EPgroup
sink()


# Investigating the effect of Close Votes, with plot in the end. See Gelman/Hill 2007 p.262
M3 <- lmer((EPGDefection == 1 & NPDefection == 0) ~ 1 + CloseVote +
  (1+CloseVote|EPgroup) +  (1|ID.EPGDefection),
          data=RegData[which(RegData$VotingRule=="s"),], family=binomial(link = "logit"))

sink("M3.txt")
summary(M3)
str(M3)
coef(M3)$EPgroup
se.coef(M3)$EPgroup
sink()

M3.pooled <- glm((EPGDefection == 1 & NPDefection == 0) ~ CloseVote,
                 data=RegData[which(RegData$VotingRule=="s"),], family=binomial(link = "logit"))
summary(M3.pooled)
M3.unpooled <- glm((EPGDefection == 1 & NPDefection == 0) ~ CloseVote + EPgroup - 1,
                data=RegData[which(RegData$VotingRule=="s"),], family=binomial(link = "logit"))
summary(M3.unpooled)

invlogit <- function(x){1/(1+exp(-x))}
display5 <- 1:5
uniq.name <- row.names(coef(M3)$EPgroup)

x <- RegData$CloseVote
y.range <- c(0,0.1)

a.hat.M3 <- coef(M3)$EPgroup[,1] #Intercept
b.hat.M3 <- coef(M3)$EPgroup[,2] #CloseVote Slope
a.se.M3 <- se.coef(M3)$EPgroup[,1]
b.se.M3 <- se.coef(M3)$EPgroup[,2]

setwd("D:/Dropbox/masteroppgave/graphs/")
pdf("closemodels.pdf", width=7, height=5)
par(mfrow=c(2,3))
for(j in display5){
  plot(0,0, type="n",
       xlim=c(-.05,1.05), bty="n", xaxp=c(0,1,1),
       ylim=y.range, xlab="Close Vote", ylab="EPr(Defection)", main=uniq.name[j])
  curve(invlogit(coef(M3.pooled)[1] + coef(M3.pooled)[2]*x), lty=2, lwd=1, add=TRUE)
  curve(invlogit(coef(M3.unpooled)[j+1] + coef(M3.unpooled)[1]*x), lwd=1, add=TRUE)
  curve(invlogit(a.hat.M3[j] + b.hat.M3[j]*x),lwd=1.5, col="red", add=TRUE)
  curve(invlogit((a.hat.M3[j]+2*a.se.M3[j]) + (b.hat.M3[j]+2*b.se.M3[j])*x),lwd=1,
        lty=2, col="red", add=TRUE)
  curve(invlogit((a.hat.M3[j]-2*a.se.M3[j]) + (b.hat.M3[j]-2*b.se.M3[j])*x),lwd=1,
        lty=2, col="red", add=TRUE)
}
plot(0,0,xlim=c(-.05,1.05), ylim=c(0,1),  type="n", axes=F, xlab="", ylab="")
legend(0,1, c("Pooled Logit", "Unpooled Logit", "Multi-level Logit"), lty=c(2,1,1),
       col=c("black", "black", "red"),lwd=c(1,1,1.5), cex=1)
dev.off()

# Same, but for abstain
```

```
MAbstain <- lmer((EPGDefection == 1 & NPDefection == 0) ~ 1 + (Voted==3) +
  (1+(Voted==3)|EPgroup) +  (1|ID.EPGDefection),
           data=RegData, family=binomial(link = "logit"))

setwd("D:/Dropbox/masteroppgave/tex/")
sink("MAbstain.txt")
summary(MAbstain)
str(MAbstain)
coef(MAbstain)
se.coef(MAbstain)
sink()

invlogit <- function(x){1/(1+exp(-x))}
jitter.binary <- function(a, jitt=.05){
  ifelse (a==0, runif(length(a), 0, jitt), runif(length(a), 1-jitt, 1))
}

display5 <- 1:5
uniq.name <- row.names(coef(MAbstain)$EPgroup)

x <- (RegData$Voted==3)
y <- (RegData$EPGDefection == 1 & RegData$NPDefection == 0)
y.jitter <- jitter.binary(y, jitt=.45)

y.range <- c(0,1)

a.hat.MAbstain <- coef(MAbstain)$EPgroup[,1] #Intercept
a.se.MAbstain <- se.coef(MAbstain)$EPgroup[,1]
b.hat.MAbstain <- coef(MAbstain)$EPgroup[,2] #Voted Abstain slope
b.se.MAbstain <- se.coef(MAbstain)$EPgroup[,2]
x.jitter <- jitter.binary(x, jitt=.45) #Jittered data for plotting

setwd("D:/Dropbox/masteroppgave/graphs/")
pdf("abstainmodels.pdf", width=7, height=5)
par(mfrow=c(2,3))
for(j in display5){
  plot(x.jitter[RegData$EPgroup==uniq.name[j]], y.jitter[RegData$EPgroup==uniq.name[j]],
       xlim=c(-.05,1.05), pch=".", col="lightblue", bty="n", xaxp=c(0,1,1),
       ylim=y.range, xlab="Voted Abstain", ylab="Pr(Defection)/Defection", main=uniq.name[j])
  curve(invlogit(a.hat.MAbstain[j] + b.hat.MAbstain[j]*x),lwd=1.5, col="red", add=TRUE)
  curve(invlogit((a.hat.MAbstain[j]+(2*a.se.MAbstain[j])) +
    (b.hat.MAbstain[j]+(2*b.se.MAbstain[j]))*x),lwd=1, col="black", lty=2, add=TRUE)
  curve(invlogit((a.hat.MAbstain[j]-(2*a.se.MAbstain[j])) +
    (b.hat.MAbstain[j]-(2*b.se.MAbstain[j]))*x),lwd=1, col="black", lty=2, add=TRUE)
}
dev.off()

#### GEE and Fixed Effects models ####

# Memory heavy, so I need to reduce the number of observations.
# I will be using Close Vote, thus only simple majority.
RegData <- RegData[which(RegData$VotingRule=="s"),]
# Use seed for replication (a randomly chosen number).
set.seed(1450)
RegData <- RegData[sample(1:dim(RegData)[1], size=100000, replace=FALSE),]
# Sort on cluster (needed for the GEE-models)
RegData <- RegData[order(RegData$ID.EPGDefection),]

#### DEFECTIONS BY NATIONAL PARTY FROM EPG #####
GEE1 <- zelig((EPGDefection == 1 & NPDefection == 0) ~ CloseVote + FinalVote + (Type=="Leg") +
  PrefDiff + EUPrefDiff + c.PdelegationSize + c.NElecRes + Gov + MemSys + DelSys + partycentered +
  c.DistrictMagnitude + EPgroup,
               data=RegData, model="logit.gee",
               id="ID.EPGDefection", corstr="exchangeable")

setwd("D:/Dropbox/masteroppgave/tex/")
sink("GEE1.txt")
summary(GEE1)
str(GEE1)
sink()
```

```
#### DEFECTION BY INDIVIDUAL MEPS ####
GEE2 <- zelig(NPDefection ~ CloseVote + FinalVote + (Type=="Leg") +
  PrefDiff + EUPrefDiff + c.PdelegationSize + c.NElecRes + Gov + MemSys + DelSys + partycentered +
  c.DistrictMagnitude + EPgroup,
               data=RegData, model="logit.gee",
               id="ID.EPGDefection", corstr="exchangeable")

sink("GEE2.txt")
summary(GEE2)
str(GEE2)
sink()

detach("package:arm")

x.out1 <- cbind(1,0,0,0,seq(0,2,0.1),0,0,0,0,0,0,0,0,0,0,0)

s.out1 <- sim(GEE1, x=x.out1)
s.out2 <- sim(GEE2, x=x.out1)

setwd("D:/Dropbox/masteroppgave/graphs/")
pdf("GEEPred.pdf", width=10, height=7)
par(mfrow=c(1,2),bty="n")
plot.zelig(s.out1, main="LR Distance, Defection by NP")
plot.zelig(s.out2, main="LR Distance, Defection from NP")
dev.off()

setwd("D:/Dropbox/masteroppgave/graphs/")
library(separationplot)
separationplot(pred=GEE1$fitted.values, file="separationplotGEE1.pdf", width=8,
               actual=GEE1$y, type="rect", line=TRUE, show.expected=TRUE,
               heading="Separation plot: GEE EPG Defection Only")
separationplot(pred=GEE2$fitted.values, file="separationplotGEE2.pdf", width=8,
               actual=GEE2$y, type="rect", line=TRUE, show.expected=TRUE,
               heading="Separation plot: GEE NP Defection")

setwd("d:/Dropbox/masteroppgave/graphs/")
pdf("rocplotGEE.pdf", width=7, height=5)
rocplot(GEE1$y, GEE2$y, fitted(GEE1), fitted(GEE2), main="GEE EPG (solid) and GEE NP (dashed)")
dev.off()


####### Fixed effects logit #######

logitModel2 <- glm((EPGDefection == 1 & NPDefection == 0) ~ PrefDiff + EUPrefDiff +
  c.PdelegationSize + c.NElecRes + Gov + MemSys + DelSys + partycentered +
  c.DistrictMagnitude + EPgroup + ID.EPGDefection,
                 data=RegData, family=binomial(link = "logit"))
summary(logitModel2)

sink("logitModel2.txt")
summary(logitModel2)
str(logitModel2)
logLik(logitModel2)
sink()

# For comparison.
logitModelSMALL <- glm((EPGDefection == 1 & NPDefection == 0) ~ PrefDiff + EUPrefDiff +
  c.PdelegationSize + c.NElecRes + Gov + MemSys + DelSys + partycentered +
  c.DistrictMagnitude + EPgroup,
                 data=RegData, family=binomial(link = "logit"))
summary(logitModelSMALL)

sink("logitModelSMALL.txt")
summary(logitModelSMALL)
str(logitModelSMALL)
logLik(logitModelSMALL)
sink()
```